MPsrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Thu Aug 21 09:14:35 1997; MasPar time 1721.16 Seconds
1336.190 Million cell updates/sec

Tabular output not generated.

Title: >US-08-320-157-6
Description: (1-2094) from US08320157.seq
Perfect Score: 2094
N.A. Sequence: 1 GAGGATCTACAGGGGACAAG.....CAAAAAAAAAACGGAGATCC 2094
Comp: CTCCTAGATGTCCCCTGTTC.....GTTTTTTTTTTGCCTCTAGG

Scoring table: TABLE default
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 362067 seqs, 549138275 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: embl-new3
1:BCT 2:FUN 3:GEN1 4:GEN2 5:HTG1 6:HTG2 7:HUM 8:INV
9:ORG 10:MAM 11:VRT 12:PLN 13:PRO 14:ROD 15:SYN 16:UNC
17:VIR

Database: genbank99
18:BCT1 19:BCT2 20:BCT3 21:BCT4 22:BCT5 23:BCT6 24:BCT7
25:BCT8 26:BCT9 27:BCT10 28:BCT11 29:GEN1 30:GEN2
31:GEN3 32:HTG1 33:HTG2 34:HTG3 35:INV1 36:INV2 37:INV3
38:INV4 39:INV5 40:INV6 41:INV7 42:INV8 43:INV9 44:INV10
45:INV11 46:MAM1 47:MAM2 48:MAM3 49:VRT1 50:VRT2 51:VRT3
52:VRT4 53:PAT1 54:PAT2 55:PAT3 56:PAT4 57:PAT5 58:PHG
59:PLN1 60:PLN2 61:PLN3 62:PLN4 63:PLN5 64:PLN6 65:PLN7
66:PLN8 67:PLN9 68:PLN10 69:PLN11 70:PRI1 71:PRI2
72:PRI3 73:PRI4 74:PRI5 75:PRI6 76:PRI7 77:PRI8 78:PRI9
79:PRI10 80:PRI11 81:PRI12 82:PRI13 83:PRI14 84:PRI15
85:ROD1 86:ROD2 87:ROD3 88:ROD4 89:ROD5 90:ROD6 91:ROD7
92:ROD8 93:STR 94:SYN 95:UNA 96:VRL1 97:VRL2 98:VRL3
99:VRL4 100:VRL5 101:VRL6 102:VRL7 103:VRL8 104:VRL9
105:VRL10

Database: genbank-new3
106:BCT 107:GEN1 108:GEN2 109:HTG1 110:HTG2 111:INV
112:MAM 113:VRT 114:PHG 115:PLN 116:PRI1 117:PRI2
118:ROD 119:SYN 120:UNA 121:VRL

Database: u-embl50_99
122:part1

Statistics: Mean 11.818; Variance 4.529; scale 2.609

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
1	2094	100.0	2094	77	HSU16811	Human Bak mRNA, compl	0.00e+00
2	1899	90.7	6478	77	HSU16812	Human Bak-2 gene, com	0.00e+00
3	1886	90.1	1949	77	HSU23765	Human Bak protein mRN	0.00e+00
4	1496	71.4	5408	77	HSU16813	Human Bak-3 pseudogen	0.00e+00
5	1274	60.8	1360	73	HSCEBP1	H.sapiens BAK mRNA fo	0.00e+00
6	1254	59.9	123579	5	HS291J10	Human DNA sequence **	0.00e+00
7	189	9.0	444	70	D88396S2	Human DNA for apotosi	7.63e-161
8	90	4.3	140	70	D88396S1	Human DNA for apotosi	3.06e-58
c 9	49	2.3	94459	86	MMTSXDNA	M.musculus 94kb genom	1.33e-19
10	38	1.8	215	57	I28278	Sequence 5 from paten	2.15e-10
c 11	33	1.6	215	57	I28278	Sequence 5 from paten	1.58e-06
12	27	1.3	1782	54	A28072	Mouse brain GAD seque	2.81e-02
13	27	1.3	45976	73	HSAC000028	00023; HTGS phase 3,	2.81e-02
c 14	25	1.2	257	77	HSU19554	Human 18q- syndrome b	5.52e-01
c 15	25	1.2	2754	79	HUMANONYMO	Human anonymous gene,	5.52e-01
16	23	1.1	105	56	I14734	Sequence 13 from pate	9.07e+00
c 17	24	1.1	105	56	I14734	Sequence 13 from pate	2.29e+00
c 18	23	1.1	259	72	HS76G6F	H.sapiens CpG island	9.07e+00
19	23	1.1	284	72	HS76G6R	H.sapiens CpG island	9.07e+00
c 20	23	1.1	421	76	HSPHKB19	H.sapiens PHKB gene (	9.07e+00
21	24	1.1	1055	85	MMGASTGN3	M.musculus gast gene	2.29e+00
c 22	23	1.1	1260	94	SYNBCTCONS	Bacterial cytochrome	9.07e+00
23	24	1.1	1267	87	MMU58136	Mus musculus gastrin	2.29e+00
c 24	23	1.1	2060	80	HUMGA16	Human G-alpha 16 prot	9.07e+00
25	22	1.1	2086	84	S52659	lactoferrin {5' regio	3.41e+01
c 26	22	1.1	2212	118	S82821S1	GSTA5=glutathione S-t	3.41e+01
c 27	22	1.1	2212	122	S82821	GSTA5=glutathione S-t	3.41e+01
28	24	1.1	2492	84	S49592	transcription factor	2.29e+00
29	24	1.1	2647	85	MMGSTP1	M.musculus (BALB/c) G	2.29e+00
30	24	1.1	2744	86	MMNMYC2	Woodchuck intronless	2.29e+00
31	24	1.1	2857	74	HSE2F1S03	Human transcription f	2.29e+00
32	23	1.1	3133	85	MMGAD67C	M.musculus mRNA for g	9.07e+00
33	23	1.1	4101	89	MUSMCKA	Mouse muscle creatine	9.07e+00
c 34	22	1.1	5374	86	MMU02313	Mus musculus MAST205	3.41e+01
c 35	24	1.1	5632	41	DROFSHB	D.melanogaster fsh me	2.29e+00
c 36	24	1.1	5931	39	DDU32174	Dictyostelium discoid	2.29e+00
37	22	1.1	6814	75	HSL219F9B	Human DNA sequence fr	3.41e+01
38	23	1.1	7096	88	MUSFLIAP	Mouse full-length int	9.07e+00
c 39	24	1.1	7792	41	DROFSHA	D.melanogaster fsh me	2.29e+00
40	24	1.1	8125	87	MMU15654	Mus musculus glutathi	2.29e+00
41	23	1.1	27371	37	CELR11F4	Caenorhabditis elegan	9.07e+00
c 42	24	1.1	35609	73	HSCU90B3	Human DNA sequence fr	2.29e+00
c 43	23	1.1	42510	78	HSV618H1	Human DNA sequence fr	9.07e+00
44	23	1.1	215285	6	HSU91322	Human chromosome 16p1	9.07e+00
45	23	1.1	215285	116	HSU91322	Human chromosome 16p1	9.07e+00

ALIGNMENTS

RESULT 1
LOCUS HSU16811 2094 bp mRNA PRI 19-AUG-1995
DEFINITION Human Bak mRNA, complete cds.
ACCESSION U16811
NID g595923
KEYWORDS .
SOURCE human.
ORGANISM Homo sapiens
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2094)
AUTHORS Kiefer,M.C., Brauer,M.J., Powers,V.C., Wu,J.J., Umansky,S.R., Tomei,L.D. and Barr,P.J.
TITLE Modulation of apoptosis by the widely distributed Bcl-2 homologue Bak
JOURNAL Nature 374 (6524), 736-739 (1995)
MEDLINE 95231654
REFERENCE 2 (bases 1 to 2094)
AUTHORS Kiefer,M.C.
TITLE Direct Submission
JOURNAL Submitted (02-NOV-1994) Michael C. Kiefer, Mol. Biol., LXR

```
          Biotechnology Inc., 1401 Marina Way South, Richmond, CA 94804, USA
FEATURES             Location/Qualifiers
     source          1..2094
                     /organism="Homo sapiens"
                     /clone="W5"
                     /clone_lib="WI-L2/ZAP II"
                     /chromosome="6"
                     /map="6p21-23"
                     /cell_line="EBV transformed B lymphoblastoid"
     CDS             201..836
                     /gene="Bak"
                     /codon_start=1
                     /product="Bak protein"
                     /db_xref="PID:g595924"
                     /translation="MASGQGPGPPRQECGEPALPSASEEQVAQDTEEVFRSYVFYRHQ
                     QEQEAEGVAAPADPEMVTLPLQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQP
                     TAENAYEYFTKIATSLFESGINWGRVVALLGFGYRLALHVYQHGLTGFLGQVTRFVVD
                     FMLHHCIARWIAQRGGWVAALNLGNGPILNVLVVLGVVLLGQFVVRRFFKS"
BASE COUNT      410 a    608 c    606 g    470 t
ORIGIN

  Query Match          100.0%;  Score 2094;  DB 77;  Length 2094;
  Best Local Similarity 100.0%;  Pred. No. 0.00e+00;
  Matches  2094;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      1 gaggatctacaggggacaagtaaaggctacatccagatgccgggaatgcactgacgcccca 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      1 GAGGATCTACAGGGGACAAGTAAAGGCTACATCCAGATGCCGGGAATGCACTGACGCCCA 60

Db     61 ttcctggaaactgggctcccactcagcccctgggagcagcagccgccagccccctcggacc 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     61 TTCCTGGAAACTGGGCTCCCACTCAGCCCCTGGGAGCAGCAGCCGCCAGCCCCTCGGACC 120

Db    121 tccatctccaccctgctgagccacccgggttgggccaggatcccggcaggctgatcccgt 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    121 TCCATCTCCACCCTGCTGAGCCACCCGGGTTGGGCCAGGATCCCGGCAGGCTGATCCCGT 180

Db    181 cctccactgagacctgaaaaatggcttcggggcaaggcccaggtcctcccaggcaggagt 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    181 CCTCCACTGAGACCTGAAAAATGGCTTCGGGGCAAGGCCCAGGTCCTCCCAGGCAGGAGT 240

Db    241 gcggagagcctgccctgccctctgcttctgaggagcaggtagcccaggacacagaggagg 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    241 GCGGAGAGCCTGCCCTGCCCTCTGCTTCTGAGGAGCAGGTAGCCCAGGACACAGAGGAGG 300

Db    301 ttttccgcagctacgtttttttaccgccatcagcaggaacaggaggctgaaggggtggctg 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    301 TTTTCCGCAGCTACGTTTTTTTACCGCCATCAGCAGGAACAGGAGGCTGAAGGGGTGGCTG 360

Db    361 cccctgccgacccagagatggtcaccttacctctgcaacctagcagcaccatggggcagg 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    361 CCCCTGCCGACCCAGAGATGGTCACCTTACCTCTGCAACCTAGCAGCACCATGGGGCAGG 420

Db    421 tgggacggcagctcgccatcatcggggacgacatcaaccgacgctatgactcagagttcc 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    421 TGGGACGGCAGCTCGCCATCATCGGGGACGACATCAACCGACGCTATGACTCAGAGTTCC 480

Db    481 agaccatgttgcagcacctgcagcccacggcagagaatgcctatgagtacttcaccaaga 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    481 AGACCATGTTGCAGCACCTGCAGCCCACGGCAGAGAATGCCTATGAGTACTTCACCAAGA 540

Db    541 ttgccaccagcctgtttgagagtggcatcaattggggccgtgtggtggctcttctgggct 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    541 TTGCCACCAGCCTGTTTGAGAGTGGCATCAATTGGGGCCGTGTGGTGGCTCTTCTGGGCT 600

Db    601 tcggctaccgtctggccctacacgtctaccagcatggcctgactggcttcctaggccagg 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    601 TCGGCTACCGTCTGGCCCTACACGTCTACCAGCATGGCCTGACTGGCTTCCTAGGCCAGG 660

Db    661 tgacccgcttcgtggtcgacttcatgctgcatcactgcattgcccggtggattgcacaga 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    661 TGACCCGCTTCGTGGTCGACTTCATGCTGCATCACTGCATTGCCCGGTGGATTGCACAGA 720

Db    721 ggggtggctgggtggcagccctgaacttgggcaatggtcccatcctgaacgtgctggtgg 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    721 GGGGTGGCTGGGTGGCAGCCCTGAACTTGGGCAATGGTCCCATCCTGAACGTGCTGGTGG 780

Db    781 ttctgggtgtggttctgttgggccagtttgtggtacgaagattcttcaaatcatgactcc 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    781 TTCTGGGTGTGGTTCTGTTGGGCCAGTTTGTGGTACGAAGATTCTTCAAATCATGACTCC 840

Db    841 caagggtgccctttgggtcccggttcagacccctgcctggacttaagcgaagtctttgcc 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    841 CAAGGGTGCCCTTTGGGTCCCGGTTCAGACCCCTGCCTGGACTTAAGCGAAGTCTTTGCC 900

Db    901 ttctctgttcccttgcagggtccccccctcaagagtacagaagctttagcaagtgtgcact 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    901 TTCTCTGTTCCCTTGCAGGGTCCCCCCTCAAGAGTACAGAAGCTTTAGCAAGTGTGCACT 960

Db    961 ccagcttcggaggccctgcgtgggggccagtcaggctgcagaggcacctcaacattgcat 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    961 CCAGCTTCGGAGGCCCTGCGTGGGGGCCAGTCAGGCTGCAGAGGCACCTCAACATTGCAT 1020

Db   1021 ggtgctagtgccctctctctgggcccagggctgtggccgtctcctccctcagctctctgg 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1021 GGTGCTAGTGCCCTCTCTCTGGGCCCAGGGCTGTGGCCGTCTCCTCCCTCAGCTCTCTGG 1080

Db   1081 gacctccttagccctgtctgctaggcgctggggagactgataacttggggaggcaagaga 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1081 GACCTCCTTAGCCCTGTCTGCTAGGCGCTGGGGAGACTGATAACTTGGGGAGGCAAGAGA 1140

Db   1141 ctgggagccacttctccccagaaagtgtttaacggttttagctttttataataccccttgt 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1141 CTGGGAGCCACTTCTCCCCAGAAAGTGTTTAACGGTTTTAGCTTTTTATAATACCCTTGT 1200

Db   1201 gagagcccattcccaccattctacctgaggccaggacgtctggggtgtggggattggtgg 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1201 GAGAGCCCATTCCCACCATTCTACCTGAGGCCAGGACGTCTGGGGTGTGGGGATTGGTGG 1260

Db   1261 gtctatgttccccaggattcagctattctggaagatcagcaccctaagagatgggactag 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1261 GTCTATGTTCCCCAGGATTCAGCTATTCTGGAAGATCAGCACCCTAAGAGATGGGACTAG 1320

Db   1321 gacctgagcctggtcctggccgtccctaagcatgtgtcccaggagcaggacctactagga 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1321 GACCTGAGCCTGGTCCTGGCCGTCCCTAAGCATGTGTCCCAGGAGCAGGACCTACTAGGA 1380

Db   1381 gagggggccaaggtcctgctcaactctacccctgctcccattcctccctccggccatac 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1381 GAGGGGGGCCAAGGTCCTGCTCAACTCTACCCCTGCTCCCATTCCTCCCTCCGGCCATAC 1440

Db   1441 tgcctttgcagttggactctcagggattctgggcttggggtgtggggtggggtggagtcg 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1441 TGCCTTTGCAGTTGGACTCTCAGGGATTCTGGGCTTGGGGTGTGGGGTGGGGTGGAGTCG 1500

Db   1501 cagaccagagctgtctgaactcacgtgtcagaagcctccaagcctgcctcccaaggtcct 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1501 CAGACCAGAGCTGTCTGAACTCACGTGTCAGAAGCCTCCAAGCCTGCCTCCCAAGGTCCT 1560

Db   1561 ctcagttctctcccttcctctctccttatagacacttgctcccaacccattcactacagg 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1561 CTCAGTTCTCTCCCTTCCTCTCTCCTTATAGACACTTGCTCCCAACCCATTCACTACAGG 1620

Db   1621 tgaaggctctcacccatccctgggggccttgggtgagtggcctgctaaggctcctccttg 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1621 TGAAGGCTCTCACCCATCCCTGGGGGCCTTGGGTGAGTGGCCTGCTAAGGCTCCTCCTTG 1680

Db   1681 cccagactacagggcttaggacttggtttgttatatcagggaaaaggagtagggagttca 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1681 CCCAGACTACAGGGCTTAGGACTTGGTTTGTTATATCAGGGAAAAGGAGTAGGGAGTTCA 1740

Db   1741 tctggagggttctaagtgggagaaggactatcaacaccactaggaatcccagaggtggat 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1741 TCTGGAGGGTTCTAAGTGGGAGAAGGACTATCAACACCACTAGGAATCCCAGAGGTGGAT 1800

Db   1801 cctccctcatggctctggcacagtgtaatccaggggtgtagatgggggaactgtgaatac 1860
```

```
Qy   1801 CCTCCCTCATGGCTCTGGCACAGTGTAATCCAGGGGTGTAGATGGGGGAACTGTGAATAC 1860

Db   1861 ttgaactctgttcccccaccctccatgctcctcacctgtctaggtctcctcagggtgggg 1920
Qy   1861 TTGAACTCTGTTCCCCCACCCTCCATGCTCCTCACCTGTCTAGGTCTCCTCAGGGTGGGG 1920

Db   1921 ggtgacagtgccttctctattggcacagcctagggtcttgggggtcaggggggagaagtt 1980
Qy   1921 GGTGACAGTGCCTTCTCTATTGGCACAGCCTAGGGTCTTGGGGGTCAGGGGGGAGAAGTT 1980

Db   1981 cttgattcagccaaatgcagggaggggaggcagatggagcccataggccacccccctatcc 2040
Qy   1981 CTTGATTCAGCCAAATGCAGGGAGGGGAGGCAGATGGAGCCCATAGGCCACCCCCTATCC 2040

Db   2041 tctgagtgtttggaaataaactgtgcaatcccctcaaaaaaaaaacggagatcc 2094
Qy   2041 TCTGAGTGTTTGGAAATAAACTGTGCAATCCCCTCAAAAAAAAAACGGAGATCC 2094


RESULT    2
LOCUS       HSU16812     6478 bp    DNA             PRI       19-AUG-1995
DEFINITION  Human Bak-2 gene, complete cds.
ACCESSION   U16812
NID         g595925
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 6478)
  AUTHORS   Kiefer,M.C., Brauer,M.J., Powers,V.C., Wu,J.J., Umansky,S.R.,
            Tomei,L.D. and Barr,P.J.
  TITLE     Modulation of apoptosis by the widely distributed Bcl-2 homologue
            Bak
  JOURNAL   Nature 374 (6524), 736-739 (1995)
  MEDLINE   95231654
REFERENCE   2  (bases 1 to 6478)
  AUTHORS   Kiefer,M.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-NOV-1994) Michael C. Kiefer, Mol. Biol., LXR
            Biotechnology Inc., 1401 Marina Way South, Richmond, CA 94804, USA
FEATURES             Location/Qualifiers
     source          1..6478
                     /organism="Homo sapiens"
                     /clone="A4"
                     /clone_lib="pWE 15/Human genomic DNA"
                     /chromosome="20"
     CDS             3279..3914
                     /gene="Bak-2"
                     /codon_start=1
                     /product="Bak-2 protein"
                     /db_xref="PID:g595926"
                     /translation="MASGQGPGPPRQECGEPALPSASEEQVAQDTEEVFRSYVFYHHQ
                     QEQEAEGAAAPADPEMVTLPLQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQP
                     TAENAYEYFTKIASSLFESGINWGRVVALLGFSYRLALHIYQRGLTGFLGQVTRFVVD
                     FMLHHCIARWIAQRGGWVAALNLGNGPILNVLVVLGVVLLGQFVVRRFFKS"
BASE COUNT     1503 a    1609 c    1599 g    1767 t
ORIGIN

  Query Match           90.7%;  Score 1899;  DB 77;  Length 6478;
  Best Local Similarity  97.3%;  Pred. No. 0.00e+00;
  Matches  2026;  Conservative    0;  Mismatches  43;  Indels  14;  Gaps  11;

Db   3090 gggacaagtaaaggctacatccagatgctgggaatgcactgacgcccattcctggaaact 3149
Qy     13 GGGACAAGTAAAGGCTACATCCAGATGCCGGGAATGCACTGACGCCCATTCCTGGAAACT 72

Db   3150 gggctcccactcagcccctgggagcagcagccgccagcccctcgggacctccatctccac 3209
Qy     73 GGGCTCCCACTCAGCCCCTGGGAGCAGCAGCCGCCAGCCCCTCGG-ACCTCCATCTCCAC 131

Db   3210 cctgctgagccacccgggttgggccaggatcccggcaggctgatcccgtcctccactgag 3269
Qy    132 CCTGCTGAGCCACCCGGGTTGGGCCAGGATCCCGGCAGGCTGATCCCGTCCTCCACTGAG 191

Db   3270 acctgaaaaatggcttcggggcaaggcccaggtcctcccaggcaggagtgcggagagcct 3329
Qy    192 ACCTGAAAAATGGCTTCGGGGCAAGGCCCAGGTCCTCCCAGGCAGGAGTGCGGAGAGCCT 251

Db   3330 gccctgccctctgcttctgaggagcaggtagcccaggacacagaggaggttttccgcagc 3389
Qy    252 GCCCTGCCCTCTGCTTCTGAGGAGCAGGTAGCCCAGGACACAGAGGAGGTTTTCCGCAGC 311

Db   3390 tacgtttttttaccaccatcagcaggaacaggaggctgaaggggcggctgcccctgccgac 3449
Qy    312 TACGTTTTTTTACCGCCATCAGCAGGAACAGGAGGCTGAAGGGGTGGCTGCCCCTGCCGAC 371

Db   3450 ccagagatggtcaccttacctctgcaacctagcagcaccatggggcaggtgggacggcag 3509
Qy    372 CCAGAGATGGTCACCTTACCTCTGCAACCTAGCAGCACCATGGGGCAGGTGGGACGGCAG 431

Db   3510 ctcgccatcattggggacgacatcaaccgacgctatgactcagagttccagaccatgttg 3569
Qy    432 CTCGCCATCATCGGGGACGACATCAACCGACGCTATGACTCAGAGTTCCAGACCATGTTG 491

Db   3570 cagcacctgcagcccacggcagagaatgcctatgagtacttcaccaagattgcctccagc 3629
Qy    492 CAGCACCTGCAGCCCACGGCAGAGAATGCCTATGAGTACTTCACCAAGATTGCCACCAGC 551

Db   3630 ctgtttgagagtggcatcaattggggccgtgtggtggctcttctgggcttcagctaccgt 3689
Qy    552 CTGTTTGAGAGTGGCATCAATTGGGGCCGTGTGGTGGCTCTTCTGGGCTTCGGCTACCGT 611

Db   3690 ctggccctacacatctaccagcgtggcctgactggcttcctgggccaggtgacccgcttt 3749
Qy    612 CTGGCCCTACACGTCTACCAGCATGGCCTGACTGGCTTCCTAGGCCAGGTGACCCGCTTC 671

Db   3750 gtggtggacttcatgctgcatcactgcattgcccggtggattgcacagaggggtggctgg 3809
Qy    672 GTGGTCGACTTCATGCTGCATCACTGCATTGCCCGGTGGATTGCACAGAGGGGTGGCTGG 731

Db   3810 gtggcagccctgaacttgggcaatggtcccatcctgaacgtgctggtggttctgggtgtg 3869
Qy    732 GTGGCAGCCCTGAACTTGGGCAATGGTCCCATCCTGAACGTGCTGGTGGTTCTGGGTGTG 791

Db   3870 gttctgttgggccagtttgtggtacgaagattcttcaaatcatgactcccaagggtgccc 3929
Qy    792 GTTCTGTTGGGCCAGTTTGTGGTACGAAGATTCTTCAAATCATGACTCCCAAGGGTGCCC 851

Db   3930 tttggggtcccagttcagaccccctgcctggacttaagcgaagtctttgccttctctgctc 3989
Qy    852 TTTGGG-TCCCGGTTCAGACCCCTGCCTGGACTTAAGCGAAGTCTTTGCCTTCTCTGTTC 910

Db   3990 cttgcaggggtccccccctcaagagtacagaagctttagcaagtgtgcactccagcttcgg 4049
Qy    911 CCTTGCAGGGTCCCCCCTCAAGAGTACAGAAGCTTTAGCAAGTGTGCACTCCAGCTTCGG 970

Db   4050 agggcccctgtgtgggggccagtcaggctgcagaggcacctcaacattccatggtgctag 4109
Qy    971 AGG--CCCTGCGTGGGGGCCAGTCAGGCTGCAGAGGCACCTCAACATTGCATGGTGCTAG 1028

Db   4110 tgggccctctctctgggcccaggggctgtggcgtctcctccctcagctctctgggacctc 4169
Qy   1029 TG--CCCTCTCTCTGGGCCCAGGGCTGTGGCCGTCTCCTCCCTCAGCTCTCTGGGACCTC 1086

Db   4170 cttagccctgtctgctaggcgctggggagactgataacttggggaggcaagagactgggga 4229
Qy   1087 CTTAGCCCTGTCTGCTAGGCGCTGGGGAGACTGATAACTTGGGGAGGCAAGAGACTGGGA 1146

Db   4230 gccacttctccccagaaagtgtttaatggttttagctttttataataccccttgtgagagc 4289
Qy   1147 GCCACTTCTCCCCAGAAAGTGTTTAACGGTTTTAGCTTTTTATAATACCCTTGTGAGAGC 1206

Db   4290 ccattcccaccattctacctgaggccaggacgtctggggtgtggggattggtgtgtctat 4349
```

```
Qy   1207 CCATTCCCACCATTCTACCTGAGGCCAGGACGTCTGGGGTGTGGGGATTGGTGGGTCTAT 1266

Db   4350 gttccccaggattcagctattctggaagatcagcaccctaagagatgggactaggacctg 4409
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1267 GTTCCCCAGGATTCAGCTATTCTGGAAGATCAGCACCCTAAGAGATGGGACTAGGACCTG 1326

Db   4410 agcctggtcctggccgtccctaagcatgtgtcccaggagcaggacctactaggagagggg 4469
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1327 AGCCTGGTCCTGGCCGTCCCTAAGCATGTGTCCCAGGAGCAGGACCTACTAGGAGAGGGG 1386

Db   4470 ggccaaggtcctgctcaactctacccctgctcccattcctccctccggccatactgccttt 4529
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1387 GGCCAAGGTCCTGCTCAACTCTACCCCTGCTCCCATTCCTCCCTCCGGCCATACTGCCTT 1446

Db   4530 tgcagttggactctcagggattctgggcttggggtgtggggtggggtggagtcg-agacc 4588
          ||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Qy   1447 TGCAGTTGGACTCTCAGGGATTCTGGGCTTGGGGTGTGGGGTGGGGTGGAGTCGCAGACC 1506

Db   4589 agagctgtctgaactcatgtgtcagaagccctccaagcctgcctcccagggtcctctcag 4648
          |||||||||||||||| |||||||||||| |||||||||||||||| ||||||||||||
Qy   1507 AGAGCTGTCTGAACTCACGTGTCAGAAGCC-TCCAAGCCTGCCTCCCAAGGTCCTCTCAG 1565

Db   4649 ttctctcccttcctctctcctttatagacacttgctcccaacccattcactacaggtgaag 4708
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1566 TTCTCTCCCTTCCTCTCTCCTTATAGACACTTGCTCCCAACCCATTCACTACAGGTGAAG 1625

Db   4709 gctcctcacccccatccctgggccttgggtgagtaacctgctaaggc--ctccttgccca 4766
          |||| ||||||    |   ||||||||||||||  ||||||||||  ||||||||||
Qy   1626 GCTC-TCACCCATCCCTGGGGGCCTTGGGTGAGTGGCCTGCTAAGGCTCCTCCTTGCCCA 1684

Db   4767 gactacagggcttaggacttggtttgttatttcagggaaaaggagtagggagttcatctg 4826
          |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Qy   1685 GACTACAGGGCTTAGGACTTGGTTTGTTATATCAGGGAAAAGGAGTAGGGAGTTCATCTG 1744

Db   4827 gagggttctaagtgggagaaggactatcaacaccactaggaatcccagaggtgggatcct 4886
          |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Qy   1745 GAGGGTTCTAAGTGGGAGAAGGACTATCAACACCACTAGGAATCCCAGAGGTGG-ATCCT 1803

Db   4887 ccctcatggctctggcacagtgtaatccaggggtgtagatgggggaactgtgaatacttg 4946
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1804 CCCTCATGGCTCTGGCACAGTGTAATCCAGGGGTGTAGATGGGGGAACTGTGAATACTTG 1863

Db   4947 aactctgttcccccaccctccatgctcctcacctgtctaggtctcctcagggtggggggt 5006
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1864 AACTCTGTTCCCCCACCCTCCATGCTCCTCACCTGTCTAGGTCTCCTCAGGGTGGGGGGT 1923

Db   5007 gagagtgccttctctattgggcacagcctagggtcttgggggtc-gggggagaagttct 5065
          || ||||||||||||||| |||||||||||||||||||||||| ||||||||||||||||
Qy   1924 GACAGTGCCTTCTCTATTGG-CACAGCCTAGGGTCTTGGGGGTCAGGGGGGAGAAGTTCT 1982

Db   5066 tgattcagccaaatgcagggaggggaggcagatggagcccataggccacctcctatcctc 5125
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1983 TGATTCAGCCAAATGCAGGGAGGGGAGGCAGATGGAGCCCATAGGCCACCCCCTATCCTC 2042

Db   5126 tgagtgtttggaaataaactgtgcaatcccctcaaaaaaataa 5168
          |||||||||||||||||||||||||||||||||||| ||
Qy   2043 TGAGTGTTTGGAAATAAACTGTGCAATCCCCTCAAAAAAAAA 2085


RESULT    3
LOCUS       HSU23765     1949 bp    mRNA            PRI       29-MAR-1996
DEFINITION  Human Bak protein mRNA, complete cds.
ACCESSION   U23765
NID         g758797
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1949)
  AUTHORS   Chittenden,T., Harrington,E.A., O'Connor,R., Flemington,C.,
            Lutz,R.J., Evan,G.I. and Guild,B.C.
  TITLE     Induction of apoptosis by the Bcl-2 homologue Bak
  JOURNAL   Nature 374 (6524), 733-736 (1995)
  MEDLINE   95231653
REFERENCE   2  (bases 1 to 1949)
  AUTHORS   Chittenden,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-MAR-1995) Thomas Chittenden, Apoptosis Technology
            Inc., 148 Sidney St., Cambridge, MA 02139, USA
FEATURES             Location/Qualifiers
     source          1..1949
                     /organism="Homo sapiens"
                     /clone_lib="Jurkat cell cDNA library (Stratagene)"
                     /cell_line="Jurkat cell"
     CDS             65..700
                     /function="induction of apoptosis"
                     /codon_start=1
                     /product="Bak protein"
                     /db_xref="PID:g758798"
                     /translation="MASGQGPGPPRQECGEPALPSASEEQVAQDTEEVFRSYVFYRHQ
                     QEQEAEGVAAPADPEMVTLPLQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQP
                     TAENAYEYFTKIATSLFESGINWGRVVALLGFGYRLALHVYQHGLTGFLGQVTRFVVD
                     FMLHHCIARWIAQRGGWVAALNLGNGPILNVLVVLGVVLLGQFVVRRFFKS"
     polyA_site      1949
BASE COUNT      368 a    557 c    575 g    449 t
ORIGIN


  Query Match           90.1%;  Score 1886;  DB 77;  Length 1949;
  Best Local Similarity  99.5%;  Pred. No. 0.00e+00;
  Matches  1940;  Conservative    0;  Mismatches   0;  Indels   9;  Gaps   7;

Db      1 tgagccacccgggttgggccaggatcccggcaggctgatcccgtcctccactgagacctg 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    137 TGAGCCACCCGGGTTGGGCCAGGATCCCGGCAGGCTGATCCCGTCCTCCACTGAGACCTG 196

Db     61 aaaaatggcttcggggcaaggcccaggtcctcccaggcaggagtgcggagagcctgccct 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    197 AAAAATGGCTTCGGGGCAAGGCCCAGGTCCTCCCAGGCAGGAGTGCGGAGAGCCTGCCCT 256

Db    121 gccctctgcttctgaggagcaggtagcccaggacacagaggaggttttccgcagctacgt 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    257 GCCCTCTGCTTCTGAGGAGCAGGTAGCCCAGGACACAGAGGAGGTTTTCCGCAGCTACGT 316

Db    181 tttttaccgccatcagcaggaacaggaggctgaaggggtggctgcccctgccgacccaga 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    317 TTTTTACCGCCATCAGCAGGAACAGGAGGCTGAAGGGGTGGCTGCCCCTGCCGACCCAGA 376

Db    241 gatggtcaccttacctctgcaacctagcagcaccatggggcaggtgggacggcagctcgc 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    377 GATGGTCACCTTACCTCTGCAACCTAGCAGCACCATGGGGCAGGTGGGACGGCAGCTCGC 436

Db    301 catcatcggggacgacatcaaccgacgctatgactcagagttccagaccatgttgcagca 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    437 CATCATCGGGGACGACATCAACCGACGCTATGACTCAGAGTTCCAGACCATGTTGCAGCA 496

Db    361 cctgcagcccacggcagagaatgcctatgagtacttcaccaagattgccaccagcctgtt 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    497 CCTGCAGCCCACGGCAGAGAATGCCTATGAGTACTTCACCAAGATTGCCACCAGCCTGTT 556

Db    421 tgagagtggcatcaattggggccgtgtggtggctcttctgggcttcggctaccgtctggc 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    557 TGAGAGTGGCATCAATTGGGGCCGTGTGGTGGCTCTTCTGGGCTTCGGCTACCGTCTGGC 616

Db    481 cctacacgtctaccagcatggcctgactggcttcctaggccaggtgacccgcttcgtggt 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    617 CCTACACGTCTACCAGCATGGCCTGACTGGCTTCCTAGGCCAGGTGACCCGCTTCGTGGT 676

Db    541 cgacttcatgctgcatcactgcattgcccggtggattgcacagaggggtggctgggtggc 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    677 CGACTTCATGCTGCATCACTGCATTGCCCGGTGGATTGCACAGAGGGGTGGCTGGGTGGC 736

Db    601 agccctgaacttgggcaatggtcccatcctgaacgtgctggtggttctgggtgtggttct 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    737 AGCCCTGAACTTGGGCAATGGTCCCATCCTGAACGTGCTGGTGGTTCTGGGTGTGGTTCT 796
```

```
Db    661 gttgggccagtttgtggtacgaagattcttcaaatcatgactcccaagggtgccctttgg 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    797 GTTGGGCCAGTTTGTGGTACGAAGATTCTTCAAATCATGACTCCCAAGGGTGCCCTTTGG 856

Db    721 gtcccggttcagacccctgcctggacttaagcgaagtctttgccttctctgttcccttgc 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    857 GTCCCGGTTCAGACCCCTGCCTGGACTTAAGCGAAGTCTTTGCCTTCTCTGTTCCCTTGC 916

Db    781 aggggtccccccctcaagagtacagaagctttagcaagtgtgcactccagcttcggagggc 840
          ||||  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    917 AGGG-TCCCCCCTCAAGAGTACAGAAGCTTTAGCAAGTGTGCACTCCAGCTTCGGAGG-- 973

Db    841 ccctgcgtgggggccagtcaggctgcagaggcacctcaacattgcatggtgctagtgggc 900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  |
Qy    974 CCCTGCGTGGGGGCCAGTCAGGCTGCAGAGGCACCTCAACATTGCATGGTGCTAGTG--C 1031

Db    901 cctctctctgggcccaggggctgtggccgtctcctccctcagctctctgggacctcctta 960
          |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Qy   1032 CCTCTCTCTGGGCCCAGGG-CTGTGGCCGTCTCCTCCCTCAGCTCTCTGGGACCTCCTTA 1090

Db    961 gccctgtctgctaggcgctggggagactgataacttggggaggcaagagactgggagcca 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1091 GCCCTGTCTGCTAGGCGCTGGGGAGACTGATAACTTGGGGAGGCAAGAGACTGGGAGCCA 1150

Db   1021 cttctccccagaaagtgtttaacggttttagctttttataataccccttgtgagagcccat 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1151 CTTCTCCCCAGAAAGTGTTTAACGGTTTTAGCTTTTTATAATACCCTTGTGAGAGCCCAT 1210

Db   1081 tcccaccattctacctgaggccaggacgtctggggtgtggggattggtgggtctatgttc 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1211 TCCCACCATTCTACCTGAGGCCAGGACGTCTGGGGTGTGGGGATTGGTGGGTCTATGTTC 1270

Db   1141 cccaggattcagctattctggaagatcagcaccctaagagatgggactaggacctgagcc 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1271 CCCAGGATTCAGCTATTCTGGAAGATCAGCACCCTAAGAGATGGGACTAGGACCTGAGCC 1330

Db   1201 tggtcctggccgtccctaagcatgtgtcccaggagcaggacctactaggagagggggggcc 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1331 TGGTCCTGGCCGTCCCTAAGCATGTGTCCCAGGAGCAGGACCTACTAGGAGAGGGGGGCC 1390

Db   1261 aaggtcctgctcaactctacccctgctcccattcctccctccggccatactgcctttgca 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1391 AAGGTCCTGCTCAACTCTACCCCTGCTCCCATTCCTCCCTCCGGCCATACTGCCTTTGCA 1450

Db   1321 gttggactctcagggattctgggcttggggtgtggggtggggtggagtcgcagaccagag 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1451 GTTGGACTCTCAGGGATTCTGGGCTTGGGGTGTGGGGTGGGGTGGAGTCGCAGACCAGAG 1510

Db   1381 ctgtctgaactcacgtgtcagaagcctccaagcctgcctcccaaggtcctctcagttctc 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1511 CTGTCTGAACTCACGTGTCAGAAGCCTCCAAGCCTGCCTCCCAAGGTCCTCTCAGTTCTC 1570

Db   1441 tcccttcctctctccttatagacacttgctcccaacccattcactacaggtgaaggctct 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1571 TCCCTTCCTCTCTCCTTATAGACACTTGCTCCCAACCCATTCACTACAGGTGAAGGCTCT 1630

Db   1501 caccccatccctgggggccttgggtgagtggcctgctaaggctcctccttgcccagacta 1560
          |||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1631 CACCC-ATCCCTGGGGGCCTTGGGTGAGTGGCCTGCTAAGGCTCCTCCTTGCCCAGACTA 1689

Db   1561 cagggcttaggacttggtttgttatatcagggaaaaggagtagggagttcatctggaggg 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1690 CAGGGCTTAGGACTTGGTTTGTTATATCAGGGAAAAGGAGTAGGGAGTTCATCTGGAGGG 1749

Db   1621 ttctaagtgggagaaggactatcaacaccactaggaatcccagaggtgggatcctccctc 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1750 TTCTAAGTGGGAGAAGGACTATCAACACCACTAGGAATCCCAGAGGTGG-ATCCTCCCTC 1808

Db   1681 atggctctggcacagtgtaatccaggggtgtagatgggggaactgtgaatacttgaactc 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1809 ATGGCTCTGGCACAGTGTAATCCAGGGGTGTAGATGGGGGAACTGTGAATACTTGAACTC 1868

Db   1741 tgttcccccaccctccatgctcctcacctgtctaggtctcctcagggtgggggggtgacag 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1869 TGTTCCCCCACCCTCCATGCTCCTCACCTGTCTAGGTCTCCTCAGGGTGGGGGGTGACAG 1928

Db   1801 tgccttctctattgggcacagcctagggtcttgggggtcaggggggagaagttcttgatt 1860
          ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1929 TGCCTTCTCTATTGG-CACAGCCTAGGGTCTTGGGGGTCAGGGGGGAGAAGTTCTTGATT 1987

Db   1861 cagccaaatgcagggaggggaggcagatggagcccataggccacccccctatcctctgagt 1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1988 CAGCCAAATGCAGGGAGGGGAGGCAGATGGAGCCCATAGGCCACCCCCTATCCTCTGAGT 2047

Db   1921 gtttggaaataaactgtgcaatcccctca 1949
          |||||||||||||||||||||||||||||
Qy   2048 GTTTGGAAATAAACTGTGCAATCCCCTCA 2076
```

RESULT 4
LOCUS HSU16813 5408 bp DNA PRI 19-AUG-1995
DEFINITION Human Bak-3 pseudogene, complete cds.
ACCESSION U16813
NID g595927
KEYWORDS .
SOURCE human.
ORGANISM Homo sapiens
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 5408)
AUTHORS Kiefer,M.C., Brauer,M.J., Powers,V.C., Wu,J.J., Umansky,S.R., Tomei,L.D. and Barr,P.J.
TITLE Modulation of apoptosis by the widely distributed Bcl-2 homologue Bak
JOURNAL Nature 374 (6524), 736-739 (1995)
MEDLINE 95231654
REFERENCE 2 (bases 1 to 5408)
AUTHORS Kiefer,M.C.
TITLE Direct Submission
JOURNAL Submitted (02-NOV-1994) Michael C. Kiefer, Mol. Biol., LXR Biotechnology Inc., 1401 Marina Way South, Richmond, CA 94804, USA

FEATURES	Location/Qualifiers
source	1..5408
	/organism="Homo sapiens"
	/clone="C5"
	/clone_lib="pWE 15/Human genomic DNA"
	/chromosome="11"
CDS	1665..1931
	/gene="Bak-3"
	/pseudo

BASE COUNT 1369 a 1384 c 1314 g 1341 t
ORIGIN

Query Match 71.4%; Score 1496; DB 77; Length 5408;
Best Local Similarity 89.6%; Pred. No. 0.00e+00;
Matches 1880; Conservative 0; Mismatches 186; Indels 33; Gaps 23;

```
Db   1476 gggacaagcaaaggctacatccagataccagaaatgcactgacgcccgttcctggaagct 1535
          |||||||| |||||||||||||||||| || | ||||||||||||||| |||||||||  ||
Qy     13 GGGACAAGTAAAGGCTACATCCAGATGCCGGGAATGCACTGACGCCCATTCCTGGAAACT 72

Db   1536 gggctcccactcagcccctgggagcagcagcctccagcccctgggaccttcaactccac 1595
          |||||||||||||||||||||||||||||||| ||||||||  || |||| || ||||||
Qy     73 GGGCTCCCACTCAGCCCCTGGGAGCAGCAGCCGCCAGCCCCTCGG-ACCTCCATCTCCAC 131

Db   1596 cctgctgacccacgcgggttgagccagcatccctggaggctgacactgtcctccactgag 1655
          ||||||||  ||||  |||||||  |||||  |||||  | |||||||  | ||||||||||||||
Qy    132 CCTGCTGAGCCACCCGGGTTGGGCCAGGATCCCGGCAGGCTGATCCCGTCCTCCACTGAG 191

Db   1656 acctgaaaaatggcatcggggcaaggcccagggcctcccaggcaggagtgcggaaagcct 1715
          ||||||||||||||  ||||||||||||||||  ||||||||||||||||||||  |||||
Qy    192 ACCTGAAAAATGGCTTCGGGGCAAGGCCCAGGTCCTCCCAGGCAGGAGTGCGGAGAGCCT 251

Db   1716 gccctgccctctgcttctgaggagcaggtagcccaggacatggagg-ggttttccgcagc 1774
          |||||||||||||||||||||||||||||||||||||||||||||  |||| ||||||||||||
Qy    252 GCCCTGCCCTCTGCTTCTGAGGAGCAGGTAGCCCAGGACACAGAGGAGGTTTTCCGCAGC 311
```

```
Db   1775 tacgtttttttaccaccatcagcaggaacaggaggctgaaggggcggccgcccctgccgac 1834
          |||||||||||||| |||||||||||||||||||||||||||||| ||| ||||||||||||
Qy    312 TACGTTTTTTACCGCCATCAGCAGGAACAGGAGGCTGAAGGGGTGGCTGCCCCTGCCGAC 371

Db   1835 ccagagatggtcaccttgcccctccaacctagcagcaccatggggcaggtgggacggcag 1894
          |||||||||||||||| || || ||||||||||||||||||||||||||||||||||||||
Qy    372 CCAGAGATGGTCACCTTACCTCTGCAACCTAGCAGCACCATGGGGCAGGTGGGACGGCAG 431

Db   1895 ctcgccatcaccagg-acgacatcaaccggcactatgacttcggagttccagaccatgct 1953
          |||||||||| | || |||||||||||||| | |||||||| | ||||||||||||||| |
Qy    432 CTCGCCATCATCGGGGACGACATCAACCGACGCTATGACT-CAGAGTTCCAGACCATGTT 490

Db   1954 gcagcacctgcagcccacggcagagaacgcctacgagtacttcaccaagatcgcctccag 2013
          |||||||||||||||||||||||||| ||||| |||||||||||||||| ||| ||||
Qy    491 GCAGCACCTGCAGCCCACGGCAGAGAATGCCTATGAGTACTTCACCAAGATTGCCACCAG 550

Db   2014 cctgtttgagagtggcatcaaccggggccgtgtggtggctctcctgggcttcggctaccg 2073
          |||||||||||||||||||  |||||||||||||||||| |||||||||||||||||
Qy    551 CCTGTTTGAGAGTGGCATCAATTGGGGCCGTGTGGTGGCTCTTCTGGGCTTCGGCTACCG 610

Db   2074 tctggtcctacatgtctaccagcacggcttgactggcttcctgggcctggtgacccgctt 2133
          ||||| |||||| |||||||||| ||| |||||||||||||| |||| ||||||||||||
Qy    611 TCTGGCCCTACACGTCTACCAGCATGGCCTGACTGGCTTCCTAGGCCAGGTGACCCGCTT 670

Db   2134 cgtggtc---ttcatgctgcaacaaggcatcgcccggtggatctcgcagaggggcggctg 2190
          |||||||   |||||||||| ||  |||| |||||||||  | |||||||| |||||
Qy    671 CGTGGTCGACTTCATGCTGCATCACTGCATTGCCCGGTGGATTGCACAGAGGGGTGGCTG 730

Db   2191 ggtggcagccctggacttgggcaatagtcccatcctgaacgtgctggtggttgtgggtgt 2250
          |||||||||||| |||||||||| |||||||||||||||||||||||||||||| |||||||
Qy    731 GGTGGCAGCCCTGAACTTGGGCAATGGTCCCATCCTGAACGTGCTGGTGGTTCTGGGTGT 790

Db   2251 ggttctgctgggccagtttgtggtaagaagattcttcaaatcatgactcccaggggtgtc 2310
          ||||||| |||||||||||||||| |||||||||||||||||||||||||||||| ||||| |
Qy    791 GGTTCTGTTGGGCCAGTTTGTGGTACGAAGATTCTTCAAATCATGACTCCCAAGGGTGCC 850

Db   2311 ctttggggtcccagctgtgacccctgcctggacttaagccaagtctttgccttccccact 2370
          ||||||| |||| | |  |||||||||||||||||||||| |||||||||||||| |   |
Qy    851 CTTTGGG-TCCCGGTTCAGACCCCTGCCTGGACTTAAGCGAAGTCTTTGCCTTCTCTGTT 909

Db   2371 cccttgcaggggtcacccttcaaaagtacagaagctctagcaagtgtgcaccccgctgc 2430
          |||||||||| || ||| |||| |||||||||||| |||||||||||||||||| || ||| |
Qy    910 CCCTTGCAGGG-TCCCCCCTCAAGAGTACAGAAGCTTTAGCAAGTGTGCACTCCAGCTTC 968

Db   2431 ggagggcccctgcgtgggggccagtcaggctgcggaggcacctcaacattgcacggtgct 2490
          |||||  |||||||||||||||||||||||||| ||||||||||||||||||||| ||||||
Qy    969 GGAGG--CCCTGCGTGGGGGCCAGTCAGGCTGCAGAGGCACCTCAACATTGCATGGTGCT 1026

Db   2491 agtgggccctctctctgggcccaggggctgtgccctcctcccttggctctctgggacctc 2550
          ||||  | ||||||  |  |   |  |  |    |||||||||  |||||||||||||||
Qy   1027 AGTGCCCTCTCTCTGGGCCCAGGGCTGTGGCCGTCTCCTCCCTCAGCTCTCTGGGACCTC 1086

Db   2551 cttagtcttgtctgctaggcgctgcagaggctgataacttggggaagcaagagactggga 2610
          |||||  | ||||||||||||||||  ||| |||||||||||||| ||||||||||||||
Qy   1087 CTTAGCCCTGTCTGCTAGGCGCTGGGGAGACTGATAACTTGGGGAGGCAAGAGACTGGGA 1146

Db   2611 gccactcctccccagtaagtgtttaacggttttagctttttataatacccttgggagagc 2670
          |||||| |||||||| ||||||||||||||||||||||||||||||||||||||||| ||||||
Qy   1147 GCCACTTCTCCCCAGAAAGTGTTTAACGGTTTTAGCTTTTTATAATACCCTTGTGAGAGC 1206

Db   2671 ccattcccaccattctacccaaggccgggatgtctggggtgtgggggttggtgggtcgta 2730
          |||||||||||||||||  ||||| ||| ||||||||||||||| |||||||||| ||
Qy   1207 CCATTCCCACCATTCTACCTGAGGCCAGGACGTCTGGGGTGTGGGGATTGGTGGGTC-TA 1265

Db   2731 acctacgtgccccaggattcagctattctggaagatcag-agcctaagagctaggacttg 2789
             |  || ||||||||||||||||||||||||||||| | |||||||| | ||||| |
Qy   1266 ---T--GTTCCCCAGGATTCAGCTATTCTGGAAGATCAGCACCCTAAGAGATGGGACTAG 1320

Db   2790 atcctggtcctgg-cc-gtccctaagcat-catgtgtcccaggagcaggactgactggga 2846
            ||||  ||||| || | || |    |  |||||||||||||||||||  ||| |||
Qy   1321 GACCTGAGCCTGGTCCTGGCCGTCCCTAAGCATGTGTCCCAGGAGCAGGACCTACTAGGA 1380

Db   2847 gaggggga-ccaaggtcctacccagctctccccgtgcccccattcctcc-tccggccatac 2904
          |||||||  |||||||||| | || |||| ||| ||| |||||||||| ||||||||||
Qy   1381 GAGGGGGGCCAAGGTCCTGCTCAACTCTACCCCTGCTCCCATTCCTCCCTCCGGCCATAC 1440

Db   2905 tgcctttgcagttggactctcagggattctgggcttggggtgtggggcggcgtggagtaa 2964
          |||||||||||||||||||||||||||||||||||||||||||||||||| || |||||||
Qy   1441 TGCCTTTGCAGTTGGACTCTCAGGGATTCTGGGCTTGGGGTGTGGGGTGGGGTGGAGTCG 1500

Db   2965 caggccagagctgtctgaacttatgtgtcagaagcctccaagcctgcctcccaaggtcct 3024
          ||| |||||||||||||||| | ||||||||||||||||||||||||||||||||||||
Qy   1501 CAGACCAGAGCTGTCTGAACTCACGTGTCAGAAGCCTCCAAGCCTGCCTCCCAAGGTCCT 1560

Db   3025 ctcagctctctccctttcctctctccttatagatacttgctcccaacccattcactacagg 3084
          ||||| |||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Qy   1561 CTCAGTTCTCTCCCTTCCTCTCTCCTTATAGACACTTGCTCCCAACCCATTCACTACAGG 1620

Db   3085 tgaaggccctcacccatccctgggggccttgggtgagtgatgcgctaaggcccctccccg 3144
          ||||||| |||||||||||||||||||||||||||||    |||||||| |||||  |
Qy   1621 TGAAGGCTCTCACCCATCCCTGGGGGCCTTGGGTGAGTGGCCTGCTAAGGCTCCTCCTTG 1680

Db   3145 cccagactacagggcttggtttagggcttggtttgttatttcagggataaggagtaggga 3204
          ||||||||||||||||     ||| |||||||||||| ||||||| ||||||||||||
Qy   1681 CCCAGACTACAGGGCTT-----AGGACTTGGTTTGTTATATCAGGGAAAAGGAGTAGGGA 1735

Db   3205 gttcatctggaaggttctaagtgggagaaggactatcaacaccac-aggaatcccagagg 3263
          |||||||||| |||||||||||||||||||||||||||||||||| ||||||||||||||
Qy   1736 GTTCATCTGGAGGGTTCTAAGTGGGAGAAGGACTATCAACACCACTAGGAATCCCAGAGG 1795

Db   3264 tgggatcctccctcatggctctggcacagtgtaatccaggggtggagataggg-aactgt 3322
          ||| |||||||||||||||||||||||||||||||||||||||||||| |||| ||||||
Qy   1796 TGG-ATCCTCCCTCATGGCTCTGGCACAGTGTAATCCAGGGGTGTAGATGGGGGAACTGT 1854

Db   3323 gaatacctgaactctgtcccccgaccctccatgctcctcacctttctgggtctctcctca 3382
          |||||| |||||||||| |||| ||||||||||||||||||| ||| ||||||  |
Qy   1855 GAATACTTGAACTCTGTTCCCCCACCCTCCATGCTCCTCACCTGTCTAGGTCTCCTCAGG 1914

Db   3383 gtgtgggggtgagagtaccttctctatcgggcacagcctagggtgttgggggtgaagggg 3442
          ||| ||||||||| ||| |||||||||| || |||||||||||||| |||||||| | ||||
Qy   1915 GTG-GGGGGTGACAGTGCCTTCTCTATTGG-CACAGCCTAGGGTCTTGGGGGTCAGGGGG 1972

Db   3443 gagaagttcttgattcagccaaatgcagggaggggaggcagaaggagcccacaggccact 3502
          |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||| ||||
Qy   1973 GAGAAGTTCTTGATTCAGCCAAATGCAGGGAGGGGAGGCAGATGGAGCCCATAGGCCACC 2032

Db   3503 ccctatcctctgagtgtttggaaataaactgtgcaatcccatcaaaaaaaaaaaggaga 3561
          |||||||||||||||||||||||||||||||||||||||||||||| |||||||||||| |||||
Qy   2033 CCCTATCCTCTGAGTGTTTGGAAATAAACTGTGCAATCCCCTCAAAAAAAAAACGGAGA 2091
```

```
RESULT    5
LOCUS       HSCEBP1      1360 bp    RNA             PRI       05-MAY-1995
DEFINITION  H.sapiens BAK mRNA for BCl-2 homologue.
ACCESSION   X84213
NID         g804984
KEYWORDS    Bcl-2 protein; CEBP-1 gene.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1360)
  AUTHORS   Farrow,S.N., White,J.H., Martinou,I., Raven,T., Pun,K.T.,
            Grinham,C.J., Martinou,J.C. and Brown,R.
  TITLE     Cloning of a bcl-2 homologue by interaction with adenovirus E1B 19K
  JOURNAL   Nature 374 (6524), 731-733 (1995)
  MEDLINE   95231652
  REMARK    Erratum:[Nature 1995 Jun 1;375(6530):431]]
REFERENCE   2  (bases 1 to 1360)
  AUTHORS   Brown,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-JAN-1995) R. Brown, Glaxo Research & Development,
            Greenford Road, Greenford, Middlesex UB6 0HE, UK
FEATURES             Location/Qualifiers
     source          1..1360
```

```
                     /organism="Homo sapiens"
                     /tissue_type="B-cell"
                     /clone_lib="EBV-transformed B-cell"
     CDS             193..828
                     /gene="BAK"
                     /codon_start=1
                     /product="bcl-2 homologue"
                     /db_xref="PID:g804985"
                     /translation="MASGQGPGPPRQECGEPALPSASEEQVAQDTEEVFRSYVFYRHQ
                     QEQEAEGVAAPADPEMVTLPLQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQP
                     TAENAYEYFTKIATSLFESGINWGRVVALLGFGYRLALHVYQHGLTGFLGQVTRFVVD
                     FMLHHCIARWIAQRGGWVAALNLGNGPILNVLVVLGVVLLGQFVVRRFFKS"
BASE COUNT      257 a    405 c    400 g    298 t
ORIGIN

  Query Match           60.8%;  Score 1274;  DB 73;  Length 1360;
  Best Local Similarity  98.2%;  Pred. No. 0.00e+00;
  Matches  1333;  Conservative    0;  Mismatches  17;  Indels   7;  Gaps   7;

Db      4 gggacaagtaaaggctacatccagatgccgggaatgcactgacgcccattcctggaaact 63
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     13 GGGACAAGTAAAGGCTACATCCAGATGCCGGGAATGCACTGACGCCCATTCCTGGAAACT 72

Db     64 gggctcccactcagcccctgggagcagcagccgccagcccctcgggacctccatctccac 123
          |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Qy     73 GGGCTCCCACTCAGCCCCTGGGAGCAGCAGCCGCCAGCCCCTCGG-ACCTCCATCTCCAC 131

Db    124 cctgctgagccacccgggttgggccaggatcccggcaggctgatcccgtcctccactgag 183
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    132 CCTGCTGAGCCACCCGGGTTGGGCCAGGATCCCGGCAGGCTGATCCCGTCCTCCACTGAG 191

Db    184 acctgaaaaatggcttcggggcaaggcccaggtcctcccaggcaggagtgcggagagcct 243
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    192 ACCTGAAAAATGGCTTCGGGGCAAGGCCCAGGTCCTCCCAGGCAGGAGTGCGGAGAGCCT 251

Db    244 gccctgccctctgcttctgaggagcaggtagcccaggacacagaggaggttttccgcagc 303
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    252 GCCCTGCCCTCTGCTTCTGAGGAGCAGGTAGCCCAGGACACAGAGGAGGTTTTCCGCAGC 311

Db    304 tacgtttttttaccgccatcagcaggaacaggaggctgaaggggtggctgcccctgccgac 363
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    312 TACGTTTTTTACCGCCATCAGCAGGAACAGGAGGCTGAAGGGGTGGCTGCCCCTGCCGAC 371

Db    364 ccagagatggtcaccttacctctgcaacctagcagcaccatggggcaggtgggacggcag 423
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    372 CCAGAGATGGTCACCTTACCTCTGCAACCTAGCAGCACCATGGGGCAGGTGGGACGGCAG 431

Db    424 ctcgccatcatcggggacgacatcaaccgacgctatgactcagagttccagaccatgttg 483
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    432 CTCGCCATCATCGGGGACGACATCAACCGACGCTATGACTCAGAGTTCCAGACCATGTTG 491

Db    484 cagcacctgcagcccacggcagagaatgcctatgagtacttcaccaagattgccaccagc 543
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    492 CAGCACCTGCAGCCCACGGCAGAGAATGCCTATGAGTACTTCACCAAGATTGCCACCAGC 551

Db    544 ctgtttgagagtggcatcaattggggccgtgtggtggctcttctgggcttcggctaccgt 603
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    552 CTGTTTGAGAGTGGCATCAATTGGGGCCGTGTGGTGGCTCTTCTGGGCTTCGGCTACCGT 611

Db    604 ctggccctacacgtctaccagcatggcctgactggcttcctaggccaggtgacccgcttc 663
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    612 CTGGCCCTACACGTCTACCAGCATGGCCTGACTGGCTTCCTAGGCCAGGTGACCCGCTTC 671

Db    664 gtggtcgacttcatgctgcatcactgcattgcccggtggattgcacagaggggtggctgg 723
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    672 GTGGTCGACTTCATGCTGCATCACTGCATTGCCCGGTGGATTGCACAGAGGGGTGGCTGG 731

Db    724 gtggcagccctgaacttgggcaatggtcccatcctgaacgtgctggtggttctgggtgtg 783
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    732 GTGGCAGCCCTGAACTTGGGCAATGGTCCCATCCTGAACGTGCTGGTGGTTCTGGGTGTG 791

Db    784 gttctgttgggccagtttgtggtacgaagattcttcaaatcatgactcccaagggtgccc 843
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy    792 GTTCTGTTGGGCCAGTTTGTGGTACGAAGATTCTTCAAATCATGACTCCCAAGGGTGCCC 851

Db    844 tttgggtcccggttcagacccctgcctggacttaagcgaagtctttgccttctctgttcc 903
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    852 TTTGGGTCCCGGTTCAGACCCCTGCCTGGACTTAAGCGAAGTCTTTGCCTTCTCTGTTCC 911

Db    904 cttgcaggggtccccccctcaagagtacagaagctttagcaagtgtgcactccagcttcgg 963
          ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    912 CTTGCAGGG-TCCCCCCTCAAGAGTACAGAAGCTTTAGCAAGTGTGCACTCCAGCTTCGG 970

Db    964 agggcccctgcgtgggggcagtcaggctgcagaggcacctcaacattgcatggtgctagt 1023
          ||| ||   | | |||| ||||||||||||||||||||||||||||||||||||||||||||
Qy    971 AGGCCCTGCGTG-GGGGCCAGTCAGGCTGCAGAGGCACCTCAACATTGCATGGTGCTAGT 1029

Db   1024 gggcccctctctctgggccaggggtgtg-ccgtctcctccctcagctctctgggacctcct 1082
          |  | ||||||  || ||||||| |||| ||||||||||||||||||||||||||||||||||
Qy   1030 GCCCTCTCTCTG-GGCCCAGGGCTGTGGCCGTCTCCTCCCTCAGCTCTCTGGGACCTCCT 1088

Db   1083 taaaccctgtcctgctaggcgctggggagactgataacttggggagacaagagactgggga 1142
          ||  ||||||| |||||||||||||||||||||||||||||||||| ||||||||||||||
Qy   1089 TAG-CCCTGTC-TGCTAGGCGCTGGGGAGACTGATAACTTGGGGAGGCAAGAGACTGGGA 1146

Db   1143 cccacttctccccaggaagtgtttaacggttttagctttttataatacccttgtgagagc 1202
           |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Qy   1147 GCCACTTCTCCCCAGAAAGTGTTTAACGGTTTTAGCTTTTTATAATACCCTTGTGAGAGC 1206

Db   1203 ccattcccaccattctacctgaggccaggacgtctggggtgtggggattggtgggtctat 1262
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1207 CCATTCCCACCATTCTACCTGAGGCCAGGACGTCTGGGGTGTGGGGATTGGTGGGTCTAT 1266

Db   1263 gttccccaggattcagctattctggaagatcagcaccctaagagatgggactaggacctg 1322
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1267 GTTCCCCAGGATTCAGCTATTCTGGAAGATCAGCACCCTAAGAGATGGGACTAGGACCTG 1326

Db   1323 atcctggtcctggccgtccctaagcatgtgtcccagg 1359
          | ||||||||||||||||||||||||||||||||||
Qy   1327 AGCCTGGTCCTGGCCGTCCCTAAGCATGTGTCCCAGG 1363


RESULT    6
ID   HS291J10   standard; DNA; HTG; 123579 BP.
AC   Z93017;
NI   e1041698
DT   19-MAR-1997 (Rel. 51, Created)
DT   19-MAR-1997 (Rel. 51, Last updated, Version 1)
DE   Human DNA sequence *** SEQUENCING IN PROGRESS *** from clone 291J10
KW   HTG; HTGS_PHASE1.
OS   Homo sapiens (human)
OC   Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
OC   Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   1-123579
RA   Beck S.;
RT   ;
RL   Submitted (06-NOV-1996) to the EMBL/GenBank/DDBJ databases.
RL   Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CB10 1SA,
RL   UK. E-mail enquires: humquery@sanger.ac.uk Clone requests:
RL   clonerequest@sanger.ac.uk
CC   IMPORTANT: This sequence is unfinished and does not necessarily
CC   represent the correct sequence. Work on the sequence is in progress
CC   and
CC   the release of this data is based on the understanding that the
CC   sequence may change as work continues. The sequence may be
CC   contaminated
CC   with foreign sequence from E.coli, yeast, vector, phage etc.
CC   Order of segments is not known; 800 n's separate segments.
CC   Unfinished sequence: dJ291J10  Contig_ID: 01371  Length: 3951 bp
CC   Unfinished sequence: dJ291J10  Contig_ID: 01622  Length: 1040 bp
CC   Unfinished sequence: dJ291J10  Contig_ID: 00751  Length: 6361 bp
CC   Unfinished sequence: dJ291J10  Contig_ID: 01422  Length: 1722 bp
CC   Unfinished sequence: dJ291J10  Contig_ID: 00105  Length: 26804 bp
CC   Unfinished sequence: dJ291J10  Contig_ID: 01527  Length: 1598 bp
CC   Unfinished sequence: dJ291J10  Contig_ID: 01231  Length: 1547 bp
```

```
CC   Unfinished sequence: dJ291J10  Contig_ID: 01408  Length: 2194 bp
CC   Unfinished sequence: dJ291J10  Contig_ID: 00133  Length: 4250 bp
CC   Unfinished sequence: dJ291J10  Contig_ID: 01390  Length: 4036 bp
CC   Unfinished sequence: dJ291J10  Contig_ID: 01463  Length: 7484 bp
CC   Unfinished sequence: dJ291J10  Contig_ID: 00255  Length: 14415 bp
CC   Unfinished sequence: dJ291J10  Contig_ID: 01327  Length: 9377 bp
CC   Unfinished sequence: dJ291J10  Contig_ID: 01441  Length: 6493 bp
CC   Unfinished sequence: dJ291J10  Contig_ID: 00931  Length: 21107 bp
FH   Key             Location/Qualifiers
FH
FT   source          1..123579
FT                   /organism="Homo sapiens"
FT                   /clone="291J10"
FT                   /chromosome="6"
SQ   Sequence 123579 BP; 28765 A; 27932 C; 27951 G; 27528 T; 11403 other;

  Query Match           59.9%;  Score 1254;  DB 5;  Length 123579;
  Best Local Similarity  98.8%;  Pred. No. 0.00e+00;
  Matches  1345;  Conservative    0;  Mismatches   1;  Indels  15;  Gaps  12;

Db 114183 ggtggcagccctgaacttgggcaatggtcccatcctgaacgtgctggtggttctgggtgt 114242
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    731 GGTGGCAGCCCTGAACTTGGGCAATGGTCCCATCCTGAACGTGCTGGTGGTTCTGGGTGT 790

Db 114243 ggttctgttgggccagtttgtggtacgaagattcttcaaatcatgactcccaagggtgcc 114302
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    791 GGTTCTGTTGGGCCAGTTTGTGGTACGAAGATTCTTCAAATCATGACTCCCAAGGGTGCC 850

Db 114303 ctttggggtcccggttcagacccctgcctggacttaagcgaagtctttgccttctctgtt 114362
          ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    851 CTTTGGG-TCCCGGTTCAGACCCCTGCCTGGACTTAAGCGAAGTCTTTGCCTTCTCTGTT 909

Db 114363 cccttgcaggggtccccccctcaagagtacagaagctttagcaagtgtgcactccagcttc 114422
          |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    910 CCCTTGCAGGG-TCCCCCCTCAAGAGTACAGAAGCTTTAGCAAGTGTGCACTCCAGCTTC 968

Db 114423 ggagggcccctgcgtgggggcaagtcanggctgcagaggcacctcaacattgcatggtgc 114482
          |||||  |||||||||||||||| ||||| ||||||||||||||||||||||||||||||||||||||||||
Qy    969 GGAGG--CCCTGCGTGGGGGCCAGTCA-GGCTGCAGAGGCACCTCAACATTGCATGGTGC 1025

Db 114483 tagtgggccctctctctgggcccaggggctgtggccgtctcctccctcagctctctgggga 114542
          |||||  |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Qy   1026 TAGTG--CCCTCTCTCTGGGCCCAGGG-CTGTGGCCGTCTCCTCCCTCAGCTCTCTGGGA 1082

Db 114543 cctccttagccctgtctgctaggcgctggggaganctgataacttggggaggcaagagac 114602
          |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Qy   1083 CCTCCTTAGCCCTGTCTGCTAGGCGCTGGGGAGA-CTGATAACTTGGGGAGGCAAGAGAC 1141

Db 114603 tgggagccacttctccccagaaagtgtttaacggttttagctttttataataccccttgtg 114662
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1142 TGGGAGCCACTTCTCCCCAGAAAGTGTTTAACGGTTTTAGCTTTTTATAATACCCTTGTG 1201

Db 114663 agagcccattcccaccattctacctgaggccaggacgtctggggtgtggggattggtggg 114722
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1202 AGAGCCCATTCCCACCATTCTACCTGAGGCCAGGACGTCTGGGGTGTGGGGATTGGTGGG 1261

Db 114723 tctatgttccccaggattcagctattctggaagatcagcaccctaagagatgggactagg 114782
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1262 TCTATGTTCCCCAGGATTCAGCTATTCTGGAAGATCAGCACCCTAAGAGATGGGACTAGG 1321

Db 114783 acctgagcctggtcctggccgtcccntaagcatgtgtcccaggagcaggacctactagga 114842
          |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1322 ACCTGAGCCTGGTCCTGGCCGTCCC-TAAGCATGTGTCCCAGGAGCAGGACCTACTAGGA 1380

Db 114843 gaggggggccaaggtcctgctcaactctacccctgctcccattcctccctccggccatac 114902
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1381 GAGGGGGGCCAAGGTCCTGCTCAACTCTACCCCTGCTCCCATTCCTCCCTCCGGCCATAC 1440

Db 114903 tgcctttgcagttggactctcagggattctgggcttggggtgtggggtggggtggagtcg 114962
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1441 TGCCTTTGCAGTTGGACTCTCAGGGATTCTGGGCTTGGGGTGTGGGGTGGGGTGGAGTCG 1500

Db 114963 cagaccagagctgtctgaactcacgtgtcagaagcctccaagcctgcctcccaaggtcct 115022
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1501 CAGACCAGAGCTGTCTGAACTCACGTGTCAGAAGCCTCCAAGCCTGCCTCCCAAGGTCCT 1560

Db 115023 ctcagttctctcccttcctctctccttatagacacttgctcccaacccattcactacagg 115082
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1561 CTCAGTTCTCTCCCTTCCTCTCTCCTTATAGACACTTGCTCCCAACCCATTCACTACAGG 1620

Db 115083 tgaaggctctcacccccatccctgggggccttgggtgagtggcctgctaaggctcctcct 115142
          |||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1621 TGAAGGCTCTCACCC--ATCCCTGGGGGCCTTGGGTGAGTGGCCTGCTAAGGCTCCTCCT 1678

Db 115143 tgcccagactacagggcttaggacttggtttgttatatcagggaaaaggagtagggagtt 115202
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1679 TGCCCAGACTACAGGGCTTAGGACTTGGTTTGTTATATCAGGGAAAAGGAGTAGGGAGTT 1738

Db 115203 catctggagggttctaagtgggagaaggactatcaacaccactaggaatcccagaggtgg 115262
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1739 CATCTGGAGGGTTCTAAGTGGGAGAAGGACTATCAACACCACTAGGAATCCCAGAGGTGG 1798

Db 115263 gatcctccctcatggctctggcacagtgtaatccaggggtgtagatgggggaactgtgaa 115322
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1799 -ATCCTCCCTCATGGCTCTGGCACAGTGTAATCCAGGGGTGTAGATGGGGGAACTGTGAA 1857

Db 115323 tacttgaactctgttcccccacccctccatgctcctcacctgtctaggtctcctcagggtg 115382
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1858 TACTTGAACTCTGTTCCCCCACCCTCCATGCTCCTCACCTGTCTAGGTCTCCTCAGGGTG 1917

Db 115383 gggggtgacagtgccttctctattgggcacagcctagggtcttgggggtcaggggggaga 115442
          |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1918 GGGGGTGACAGTGCCTTCTCTATTGG-CACAGCCTAGGGTCTTGGGGGTCAGGGGGGAGA 1976

Db 115443 agttcttgattcagccaaatgcagggaggggaggcagatgngagcccataggccaccccc 115502
          |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Qy   1977 AGTTCTTGATTCAGCCAAATGCAGGGAGGGGAGGCAGATG-GAGCCCATAGGCCACCCCC 2035

Db 115503 tatcctctgagtgtttggaaataaactgtgcaatcccctca 115543
          |||||||||||||||||||||||||||||||||||||||||
Qy   2036 TATCCTCTGAGTGTTTGGAAATAAACTGTGCAATCCCCTCA 2076


RESULT    7
LOCUS       D88396S2      444 bp    DNA             PRI       10-JAN-1997
DEFINITION  Human DNA for apotosis-regulator Bak, exon 2, 3 and partial cds.
ACCESSION   D88397
NID         g1655492
KEYWORDS    BaK; apotosis-regulator.
SEGMENT     2 of 2
SOURCE      Homo sapiens DNA, clone:pGEMBak01-02.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 444)
  AUTHORS   Eguchi,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-OCT-1996) to the DDBJ/EMBL/GenBank databases.
            Hidetaka Eguchi, Saitama Cancer Center Research Institute,
            Depertment of Biochemistry; 818 Komuro, Ina, Kita-adachi-gun,
            Saitama 362, Japan (E-mail:hide@saitama-cc.go.jp,
            Tel:048-722-1111(ex.255), Fax:048-722-1739)
REFERENCE   2  (bases 1 to 444)
  AUTHORS   Eguchi,H. and Hayashi,S.
  TITLE     Estrogen alters expression of apoptosis-regulators, Bcl-2, Bcl-xL
            and Bak, as well as susceptibility to therapeutic agents of human
            breast cancer cells
  JOURNAL   Unpublished (1996)
FEATURES             Location/Qualifiers
     source          1..444
                     /organism="Homo sapiens"
                     /clone="pGEMBak01-02"
     CDS             join(D88396:26..90,51..231,358..>444)
                     /codon_start=1
                     /product="BaK"
                     /db_xref="PID:g1655494"
```

```
                  /translation="MLQHLQPTAENAYEYFTKIATSLFESGINWGRVVALLGFGYRLA
                  LHVYQHGLTGFLGQVTRFVVDFMLHHCIARWIAQRGGWVAALNLGNGPILNVLVVLGV
                  VLLGQFVVR"
     intron       1..50
                  /number=1
     exon         51..231
                  /number=2
     intron       232..357
                  /number=2
     exon         358..444
                  /number=3
BASE COUNT       58 a    133 c    132 g    119 t      2 others
ORIGIN

  Query Match            9.0%;  Score 189;  DB 70;  Length 444;
  Best Local Similarity  99.0%;  Pred. No. 7.63e-161;
  Matches   191;  Conservative    0;  Mismatches   2;  Indels   0;  Gaps   0;

Db     42 tgcccacagcctgtttgagagtggcatcaattggggccgtgtggtggctctttctgggctt 101
          ||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    542 TGCCACCAGCCTGTTTGAGAGTGGCATCAATTGGGGCCGTGTGGTGGCTCTTCTGGGCTT 601

Db    102 cggctaccgtctggccctacacgtctaccagcatggcctgactggcttcctaggccaggt 161
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    602 CGGCTACCGTCTGGCCCTACACGTCTACCAGCATGGCCTGACTGGCTTCCTAGGCCAGGT 661

Db    162 gacccgcttcgtggtcgacttcatgctgcatcactgcattgcccggtggattgcacagag 221
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    662 GACCCGCTTCGTGGTCGACTTCATGCTGCATCACTGCATTGCCCGGTGGATTGCACAGAG 721

Db    222 gggtggctgggtg 234
          |||||||||||||
Qy    722 GGGTGGCTGGGTG 734


RESULT    8
LOCUS       D88396S1      140 bp    DNA             PRI       05-DEC-1996
DEFINITION  Human DNA for apotosis regulator Bak, exon 1.
ACCESSION   D88396
NID         g1655491
KEYWORDS    apotosis-regulator; Bak.
SEGMENT     1 of 2
SOURCE      Homo sapiens DNA, clone:pGEMBak01-02.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 140)
  AUTHORS   Eguchi,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-OCT-1996) to the DDBJ/EMBL/GenBank databases.
            Hidetaka Eguchi, Saitama Cancer Center Research Institute,
            Depertment of Biochemistry; 818 Komuro, Ina, Kita-adachi-gun,
            Saitama 362, Japan (E-mail:hide@saitama-cc.go.jp,
            Tel:048-722-1111(ex.255), Fax:048-722-1739)
REFERENCE   2  (bases 1 to 140)
  AUTHORS   Eguchi,H. and Hayashi,S.
  TITLE     Estrogen alters expression of apoptosis-regulators, Bcl-2, Bcl-xL
            and Bak, as well as susceptibility to therapeutic agents of human
            breast cancer cells
  JOURNAL   Unpublished (1996)
FEATURES             Location/Qualifiers
     source          1..140
                     /organism="Homo sapiens"
                     /clone="pGEMBak01-02"
     exon            1..90
                     /number=1
     intron          91..140
                     /number=1
BASE COUNT       33 a     52 c     27 g     28 t
ORIGIN

  Query Match            4.3%;  Score 90;  DB 70;  Length 140;
  Best Local Similarity 100.0%;  Pred. No. 3.06e-58;
  Matches    90;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      1 acgctatgactcagagttccagaccatgttgcagcacctgcagcccacggcagagaatgc 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    461 ACGCTATGACTCAGAGTTCCAGACCATGTTGCAGCACCTGCAGCCCACGGCAGAGAATGC 520

Db     61 ctatgagtacttcaccaagattgccaccag 90
          |||||||||||||||||||||||||||||
Qy    521 CTATGAGTACTTCACCAAGATTGCCACCAG 550


RESULT    9
LOCUS       MMTSXDNA    94459 bp    DNA             ROD       15-NOV-1996
DEFINITION  M.musculus 94kb genomic sequence encoding Tsx gene.
ACCESSION   X99946
NID         g1673440
KEYWORDS    testis-specific; Tsx gene.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Mus.
REFERENCE   1  (bases 1 to 94459)
  AUTHORS   Simmler,M.C., Cunningham,D.B., Clerc,P., Vermat,T., Caudron,D.,
            Cruaud,C., Pawlak,A., Szpirer,C., Weissenbach,J., Claverie,J.M. and
            Avner,P.
  TITLE     A 94kb genomic sequence 3' to the murine Xist gene reveals an AT
            rich region containing a new testis specific gene Tsx
  JOURNAL   Hum. Mol. Genet. In press
REFERENCE   2  (bases 1 to 94459)
  AUTHORS   Claverie,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-AUG-1996) J. Claverie, Information Genetique &
            Structurale, CNRS EP 91, 31 Chemin Joseph Aiguier, 13402 Marseille
            Cedex 20, FRANCE
COMMENT     This gene constitutes one of the only well verified examples of a
            case where the AG-GT rule in not obeyed.
FEATURES             Location/Qualifiers
     source          1..94459
                     /organism="Mus musculus"
                     /chromosome="X"
                     /clone_lib="YAC pa-2"
                     /map="D-E mu"
     mRNA            join(44627..44791,46180..46215,46896..46932,48355..48436,
                     50663..50763,52149..52207,54425..54741)
                     /gene="Tsx"
     exon            44627..44791
                     /gene="Tsx"
                     /number=1
     CDS             join(44698..44791,46180..46215,46896..46932,48355..48436,
                     50663..50763,52149..52207,54425..54450)
                     /gene="Tsx"
                     /codon_start=1
                     /db_xref="PID:e258462"
                     /db_xref="PID:g1673441"
                     /translation="MSEKQSPKTSEAECSAMDLPEFEDEENWLFKVLGFQPGPSSALD
                     DDTDDQADEPLSAAEFLHLQDILQEDRVSSTDDEDTCQAGCTEDDETSHSDRDIDNNV
                     KVITGNIKASPSMYMEMFTDQNPQADQDLEETESDGAMNPTD"
     intron          44792..46179
                     /gene="Tsx"
                     /number=1
     exon            46180..46215
                     /gene="Tsx"
                     /number=2
     intron          46216..46895
                     /gene="Tsx"
                     /number=2
     exon            46896..46932
                     /gene="Tsx"
                     /number=3
     intron          46933..48354
                     /gene="Tsx"
                     /number=3
```

```
     exon            48355..48436
                     /gene="Tsx"
                     /number=4
     intron          48437..50662
                     /gene="Tsx"
                     /number=4
     exon            50663..50763
                     /gene="Tsx"
                     /number=5
     intron          50764..52148
                     /gene="Tsx"
                     /number=5
     exon            52149..52207
                     /gene="Tsx"
                     /number=6
     intron          52208..54424
                     /gene="Tsx"
                     /number=6
     exon            54425..54741
                     /gene="Tsx"
                     /number=7
BASE COUNT    26203 a  19918 c  19582 g  28755 t      1 others
ORIGIN

  Query Match           2.3%;  Score 49;  DB 86;  Length 94459;
  Best Local Similarity  67.7%;  Pred. No. 1.33e-19;
  Matches   113;  Conservative    0;  Mismatches  52;  Indels   2;  Gaps   2;

Db  52745 tcccactcagacatccccagataagcagccc-ctcctctccccttttgttaacaaaccaag 52803
          ||||||| |||   | |||||| | |  ||  ||||| | ||||    ||||||||||||
Cp   1761 TCCCACTTAGAACCCTCCAGATGAACTCCCTACTCCTTTTCCCTGATATAACAAACCAAG 1702

Db  52804 tccaaggctcttgaatctggccaaggaggagccccagcaggctggttcagctgagatccc 52863
          ||| | || ||  | ||||| ||||||||||||  |||||||   | || |  ||  |||
Cp   1701 TCCTAAGCCCTGTAGTCTGGGCAAGGAGGAGCCTTAGCAGGCCACT-CACCCAAGGCCCC 1643

Db  52864 tggggagtagataaggttttcacctgtaggaaccagattgtgagtaa 52910
            ||||   |  |  |  ||||||||||  |   | ||| ||| ||
Cp   1642 CAGGGATGGGTGAGAGCCTTCACCTGTAGTGAATGGGTTGGGAGCAA 1596


RESULT   10
LOCUS       I28278        215 bp    DNA             PAT       30-OCT-1996
DEFINITION  Sequence 5 from patent US 5569830.
ACCESSION   I28278
NID         g1819054
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 215)
  AUTHORS   Bennett,A., Labavitch,J.M., Powell,A. and Stotz,H.
  TITLE     Plant inhibitors of fungal polygalacturonases and their use to
            control fungal disease
  JOURNAL   Patent: US 5569830-A 5 29-OCT-1996;

FEATURES             Location/Qualifiers
     source          1..215
                     /organism="unknown"
BASE COUNT       15 a      8 c     25 g     26 t    141 others
ORIGIN

  Query Match           1.8%;  Score 38;  DB 57;  Length 215;
  Best Local Similarity  16.9%;  Pred. No. 2.15e-10;
  Matches    27;  Conservative   67;  Mismatches  64;  Indels   2;  Gaps   2;

Db      2 tnvtmssssvvsrtascndkakkdgnttsswttdccnrtwgvcdtdttyrvnndsghnky 61
          | :|::: ::: : |:   :  ::|  |::  |:||   :|: : :   ::  ::|: ::
Qy    224 TCCTCCCAGGCAGGAGTGCGGAGAGCCTGCCCTGCCCTCTGCTTCTGAGGAGCAGGTAGC 283

Db     62 ssanynyggnnvgaakthyythtnvsgad-sktvtdsynasgtsssnggtdgnrsgadsy 120
          ::|   :    :| |: :::|:  : | : ::| |: :  :| :    | :| :  |::
Qy    284 CCAGGACACAGAGGAGGTTTTCCGCAGCTACGTTTTTTACCGCCATCAGCAGGAACAGGA 343

Db    121 gssktamtsrnrtgktannavdsrnmgdasvgsdkntkkh 160
          |::: |: :: :||:     :  : :|: :: ::: |:::
Qy    344 GGCTGAA-GGGGTGGCTGCCCCTGCCGACCCAGAGATGGT 382


RESULT   11
LOCUS       I28278        215 bp    DNA             PAT       30-OCT-1996
DEFINITION  Sequence 5 from patent US 5569830.
ACCESSION   I28278
NID         g1819054
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 215)
  AUTHORS   Bennett,A., Labavitch,J.M., Powell,A. and Stotz,H.
  TITLE     Plant inhibitors of fungal polygalacturonases and their use to
            control fungal disease
  JOURNAL   Patent: US 5569830-A 5 29-OCT-1996;

FEATURES             Location/Qualifiers
     source          1..215
                     /organism="unknown"
BASE COUNT       15 a      8 c     25 g     26 t    141 others
ORIGIN

  Query Match           1.6%;  Score 33;  DB 57;  Length 215;
  Best Local Similarity  13.7%;  Pred. No. 1.58e-06;
  Matches    26;  Conservative   74;  Mismatches  89;  Indels   1;  Gaps   1;

Db      9 ssvvsrtascndkakkdgnttsswttdccnrtwgvcdtdttyrvnndsghnkyssanyny 68
          ::::::| :  :: :::| | ::: |:|    :|:|: :  : :  :: : : ::  :
Cp   1025 GCACCATGCAATGTTGAGGTGCCTCTGCAGCCTGACTGGCCCCCACGC-AGGGCCTCCGA 967

Db     69 ggnnvgaakthyythtnvsgadsktvtdsynasgtsssnggtdgnrsgadsygssktamt 128
           |  :|| : : : :| :: |: :  |    |:    : || :| ::  :::  :: |:
Cp    966 AGCTGGAGTGCACACTTGCTAAAGCTTCTGTACTCTTGAGGGGGGACCCTGCAAGGGAAC 907

Db    129 srnrtgktannavdsrnmgdasvgsdkntkkhaknsadgkvgsknngdrnnrygtgtksn 188
          : :  :    |:::  :|   : ::   :::|: : : ::     |::      | ::
Cp    906 AGAGAAGGCAAAGACTTCGCTTAAGTCCAGGCAGGGGTCTGAACCGGGACCCAAAGGGCA 847

Db    189 vsnncgggnk 198
          ::   ||| :
Cp    846 CCCTTGGGAG 837


RESULT   12
LOCUS       A28072       1782 bp    RNA             PAT       11-OCT-1995
DEFINITION  Mouse brain GAD sequence (MBGAD).
ACCESSION   A28072
NID         g1247776
KEYWORDS    .
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae;
            Murinae; Mus.
REFERENCE   1  (bases 1 to 1782)
  AUTHORS   .
  TITLE     A METHOD FOR THE DIAGNOSIS AND TREATMENT OF GLUTAMIC ACID
            DECARBOXYLASE AUTOANTIGEN ASSOCIATED DISEASES
  JOURNAL   Patent: WO 9214485-A 11 03-SEP-1992;

FEATURES             Location/Qualifiers
     source          1..1782
                     /organism="Mus musculus"
     CDS             1..1782
                     /gene="GAD"
                     /codon_start=1
                     /db_xref="PID:e204340"
                     /db_xref="PID:g1247777"
```

```
                     /translation="MASSTPSPATSSNAGADPNTTNLRPTTYDTWCGVAHGCTRKLGL
                     KICGFLQRTNSLEEKSRLVSAFRERQSSKNLLSCENSDQGARFRRTETDFSNLFAQDL
                     LPAKNGEEQTAQFLLEVVDILLNYVRKTFDRSTKVLDFHHPHQLLEGMEGFNLELSDH
                     PESLEQILVDCRDTLKYGVRTGHPRFFNQLSTGLDIIGLAGEWLTSTANTNMFTYEIA
                     PVFVLMEQITLKKMRKIVGWSNKDGDGIFSPGGAISNMYSIMAARYKYFPEVKTKGMA
                     AVPKLVLFTSEHSHYSIKKAGAALGFGTDNVILIKCNERGKIIPADLEAKILDAKQKG
                     YVPLYVNATAGTTVYGAFDPIQEIADICEKYNLWLHVDAAWGGGLLMSRKHRHKLSGI
                     ERANSVTWNPHKMMGVLLQCSAILVKEKGILQGCNQMCAGYLFQPDKQYDVSYDTGDK
                     AIQCGRHVDIFKFWLMWKAKGTVGFENQINKCLELADYLYAKIKNREEFEMVFDGEPE
                     HTNVCFWYIPQSLRGVPDSPERREKLHRVAPKIKALMMESGTTMVGYQPQGDKANFFR
                     MVISNPAATQSDIDFLIEEIERLGQDL"
BASE COUNT      474 a    440 c    458 g    410 t
ORIGIN

  Query Match           1.3%;  Score 27;  DB 54;  Length 1782;
  Best Local Similarity  71.4%;  Pred. No. 2.81e-02;
  Matches    45;  Conservative    0;  Mismatches  18;  Indels   0;  Gaps   0;

Db   1715 cagccgccacccagtctgacatcgatttcctcattgaggagatagagaggttgggccagg 1774
          ||||||||| ||  || ||| || || ||| |  ||  |||  |    ||||||||||||
Qy    100 CAGCCGCCAGCCCCTCGGACCTCCATCTCCACCCTGCTGAGCCACCCGGGTTGGGCCAGG 159

Db   1775 atc 1777
          |||
Qy    160 ATC 162


RESULT   13
LOCUS       HSAC000028  45976 bp    DNA             PRI       21-OCT-1996
DEFINITION  00023; HTGS phase 3, complete sequence.
ACCESSION   AC000028
NID         g1620534
KEYWORDS    HTG; HTGS_PHASE3.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 45976)
  AUTHORS   Wang,Z. and Roe,B.A.
  TITLE     completed clone 6e4 from human chromosome 12, 45976 bp
  JOURNAL   Unpublished (1996)
REFERENCE   2  (bases 1 to 45976)
  AUTHORS   Meltzer,P.
  JOURNAL   Unpublished (1996)
REFERENCE   3  (bases 1 to 45976)
  AUTHORS   Roe,B.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-OCT-1996) Department Of Chemistry And Biochemistry,
            The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
            OK 73019, USA
FEATURES             Location/Qualifiers
     source          1..45976
                     /organism="Homo sapiens"
                     /chromosome="12"
BASE COUNT    10870 a  12130 c  12219 g  10524 t    233 others
ORIGIN

  Query Match           1.3%;  Score 27;  DB 73;  Length 45976;
  Best Local Similarity  80.0%;  Pred. No. 2.81e-02;
  Matches    36;  Conservative    0;  Mismatches   9;  Indels   0;  Gaps   0;

Db  21645 cctcccacccacctgcactctcagctctctcccttcacctctcctt 21689
          ||| || ||||  |  ||||||| ||||||||||  ||||||||
Qy   1543 CCTGCCTCCCAAGGTCCTCTCAGTTCTCTCCCTTCCTCTCTCCTT 1587


RESULT   14
LOCUS       HSU19554      257 bp    DNA             PRI       02-DEC-1995
DEFINITION  Human 18q- syndrome breakpoint satellite III DNA recombination
            sequence.
ACCESSION   U19554
NID         g1098521
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 257)
  AUTHORS   Schneider,S.S., Trask,B.J., Overhauser,J. and Silverman,G.A.
  TITLE     An 18q- syndrome breakpoint is located in a cluster of serpin genes
            and was formed by an illegitimate recombination with satellite III
            DNA
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 257)
  AUTHORS   Silverman,G.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-JAN-1995) Gary A. Silverman, Pediatrics, Harvard
            Medical School, Joint Program in Neonatology, 300 Longwood Avenue
            Enders-9, Boston, MA 02115, USA
FEATURES             Location/Qualifiers
     source          1..257
                     /organism="Homo sapiens"
                     /isolate="JL184. an 18q syndrome patient"
                     /map="18q21.3"
                     /chromosome="18"
     misc_feature    <1..45
                     /note="normal chromosome 18 DNA; centromeric"
     misc_recomb     46..79
                     /note="unknown filler sequence"
                     /organism="Homo sapiens"
     satellite       80..256
                     /rpt_family="satellite III"
BASE COUNT       77 a     27 c     90 g     63 t
ORIGIN

  Query Match           1.2%;  Score 25;  DB 77;  Length 257;
  Best Local Similarity  71.2%;  Pred. No. 5.52e-01;
  Matches    42;  Conservative    0;  Mismatches  17;  Indels   0;  Gaps   0;

Db    179 agtgaatggattggggtggagtgtataagagtggagaagagtggagtgaaatgaaggga 237
          |||||||||  |||||    |||||  ||  ||  ||||  ||  ||||  |||  |||  |||
Cp   1615 AGTGAATGGGTTGGGAGCAAGTGTCTATAAGGAGAGAGGAAGGGAGAGAACTGAGAGGA 1557


RESULT   15
LOCUS       HUMANONYMO   2754 bp    DNA             PRI       31-DEC-1994
DEFINITION  Human anonymous gene, complete cds.
ACCESSION   L18972
NID         g388011
KEYWORDS    .
SOURCE      Homo sapiens (tissue library: Stratagene #936206) fetus DNA.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2754)
  AUTHORS   Xie,Y.G., Han,F.Y., Peyrard,M., Ruttledge,M.H., Fransson,I.,
            DeJong,P., Collins,J., Dunham,I., Nordenskjold,M. and Dumanski,J.P.
  TITLE     Cloning of a novel, anonymous gene from a megabase-range YAC and
            cosmid contig in the neurofibromatosis type 2/meningioma region on
            human chromosome 22q12
  JOURNAL   Hum. Mol. Genet. 2 (9), 1361-1368 (1993)
  MEDLINE   94061029
FEATURES             Location/Qualifiers
     source          1..2754
                     /organism="Homo sapiens"
                     /dev_stage="fetus"
                     /tissue_lib="Stratagene #936206"
                     /map="22q12"
     CDS             35..2086
                     /gene="anonymous"
                     /codon_start=1
                     /db_xref="PID:g388012"
                     /translation="MSSESSKKRKPKVIRSDGAPAEGKRNRSDTEQEGKYYSEEAEVD
                     LRDPGRDYELYKYTCQELQRLMAEIQDLKSRGGKDVAIEIEERRIQSCVHFMTLKKLN
                     RLAHIRLKKGRDQTHEAKQKVDAYHLQLQNLLYEVMHLQKEITKCLEFKSKHEEIDLV
                     SLEEFYKEAPPDISKAEVTMGDPHQQTLARLDWELEQRKRLAEKYRECLSNKEKILKE
```

IEVKKEYLSSLQPRLNSIMQASLPVQEYLFMPFDQAHKQYETARHLPPPLYVLFVQAT
AYGQACDKTLSVAIEGSVDEAKALFKPPEDSQDDESDSDAEEEQTTKRRRPTLGVQLD
DKRKEMLKRHPLSVMLDLKCKDDSVLHLTFYYLMNLNIMTVKAKVTTAMELITPISAG
DLLSPDSVLSCLYPGDHGKKTPNPANQYQFDKVGILTLSDYVLELGHPYLWVQKLGGL
HFPKEQPQQTVIADHSLSASHMETTMKLLKTRVQSRLALHKQFASLEHGIVPVTSDCQ
YLFPAKVVSRLVKWVTIAHEDYMELHFTKDIVDAGLAGDTNLYYMALIERGTAKLQAA
VVLNPGYSSIPPIFQLCLNWKGEKTNSNDDNIRAMEGEVNVCYKELCGPWPSHQLLTN
QLQRLCVLLDVYLETESHDDSVEGPKEFPQEKMCLRLFRGPSRMKPFKYNHPQGFFSH
R"

BASE COUNT 744 a 694 c 723 g 593 t
ORIGIN

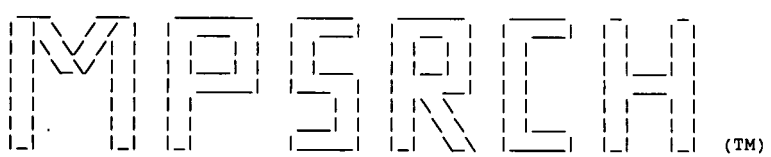
Query Match 1.2%; Score 25; DB 79; Length 2754;
Best Local Similarity 68.1%; Pred. No. 5.52e-01;
Matches 47; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

Db 229 gatggctgagatccaagacctgaagagcaggggtggcaaggatgtggcaatagaaataga 288
||||| ||||| || ||||| || | | |||| | || | ||| | ||| | |||
Cp 1638 GATGGGTGAGAGCCTTCACCTGTAGTGAATGGGTTGGGAGCAAGTGTCTATAAGGAGAGA 1579

Db 289 agaacggag 297
||| ||||
Cp 1578 GGAAGGGAG 1570

Search completed: Thu Aug 21 10:27:37 1997
Job time : 4382 secs.

****



****

Release 2.1D John F. Collins, Biocomputing Research Unit.

MPsrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Thu Aug 21 10:27:57 1997; MasPar time 218.06 Seconds
944.863 Million cell updates/sec

Tabular output not generated.

Title: >US-08-320-157-6
Description: (1-2094) from US08320157.seq
Perfect Score: 2094
N.A. Sequence: 1 GAGGATCTACAGGGGACAAG.....CAAAAAAAAAACGGAGATCC 2094
Comp: CTCCTAGATGTCCCCTGTTC.....GTTTTTTTTTTGCCTCTAGG

Scoring table: TABLE default
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 134151 seqs, 49196315 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: n-geneseq27
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20 21:part21 22:part22 23:part23
24:part24 25:part25 26:part26 27:part27

Statistics: Mean 9.693; Variance 5.444; scale 1.780

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description	Pred. No.
	1	2094	100.0	2094	24	T42138	Bak gene.	0.00e+00
	2	2072	98.9	2072	14	Q95492	Human Cdn-1 cDNA.	0.00e+00
	3	1899	90.7	6511	14	Q95493	Human Cdn-2 DNA.	0.00e+00
	4	1886	90.1	1968	19	T17375	Bcl-Y cDNA.	0.00e+00
	5	1496	71.4	5408	14	Q95494	Human Cdn-3 DNA.	0.00e+00
	6	883	42.2	1286	24	T42139	Bak-2 gene.	0.00e+00
c	7	95	4.5	1047	2	Q10572	Human Natriuretic Pep	4.06e-46
	8	78	3.7	1047	2	Q10572	Human Natriuretic Pep	2.35e-34
	9	43	2.1	91	9	Q51746	Oligonucleotide probe	1.89e-11
	10	44	2.1	204	1	N81164	Base substituted E.co	4.60e-12
c	11	45	2.1	204	1	N81164	Base substituted E.co	1.11e-12
c	12	42	2.0	91	9	Q51746	Oligonucleotide probe	7.69e-11
c	13	37	1.8	114	12	Q70465	Generic DNA sequence	7.46e-08
c	14	36	1.7	114	12	Q70468	Generic DNA sequence	2.86e-07
	15	34	1.6	114	12	Q70466	Generic DNA sequence	4.05e-06
	16	34	1.6	114	12	Q70465	Generic DNA sequence	4.05e-06
	17	34	1.6	114	12	Q70469	Generic DNA sequence	4.05e-06
	18	34	1.6	114	12	Q70467	Generic DNA sequence	4.05e-06
	19	34	1.6	114	12	Q70468	Generic DNA sequence	4.05e-06
c	20	34	1.6	114	12	Q70467	Generic DNA sequence	4.05e-06
c	21	34	1.6	114	12	Q70466	Generic DNA sequence	4.05e-06
c	22	33	1.6	114	12	Q70469	Generic DNA sequence	1.49e-05
	23	34	1.6	3871	2	N71302	HSV-1 gB and surround	4.05e-06
	24	31	1.5	114	12	Q70470	Generic DNA sequence	1.95e-04
c	25	32	1.5	114	12	Q70470	Generic DNA sequence	5.44e-05
	26	29	1.4	114	12	Q70472	Generic DNA sequence	2.38e-03
c	27	30	1.4	114	12	Q70471	Generic DNA sequence	6.87e-04
c	28	29	1.4	114	12	Q70472	Generic DNA sequence	2.38e-03
	29	28	1.3	114	12	Q70473	Generic DNA sequence	8.11e-03
c	30	27	1.3	114	12	Q70473	Generic DNA sequence	2.71e-02
	31	27	1.3	1782	5	Q28435	Brain GAD #2.	2.71e-02
c	32	26	1.2	36	2	Q11195	Ballast Constituent c	8.86e-02
c	33	26	1.2	39	7	Q51787	Mixed oligonucleotide	8.86e-02
c	34	25	1.2	74	21	T13613	DC43 TSAR library gen	2.83e-01
	35	25	1.2	75	21	T13612	DC43 TSAR library gen	2.83e-01
c	36	25	1.2	81	21	T13611	DC43 TSAR library gen	2.83e-01
	37	25	1.2	82	21	T13610	DC43 TSAR library gen	2.83e-01
	38	26	1.2	114	12	Q70471	Generic DNA sequence	8.86e-02
	39	26	1.2	565	6	Q35072	HCV envelope region n	8.86e-02
c	40	25	1.2	565	6	Q35072	HCV envelope region n	2.83e-01
c	41	23	1.1	66	21	T13585	TSAR-9 library genera	2.69e+00
c	42	24	1.1	75	21	T13612	DC43 TSAR library gen	8.85e-01
c	43	24	1.1	82	21	T13610	DC43 TSAR library gen	8.85e-01
	44	24	1.1	2456	9	Q51137	Retinoblastoma-associ	8.85e-01
	45	24	1.1	2492	10	Q67603	Retinoblastoma protei	8.85e-01

ALIGNMENTS

RESULT 1
ID T42138 standard; cDNA; 2094 BP.
AC T42138;
DT 22-FEB-1997 (first entry)
DE Bak gene.
KW Human; Bak; apoptosis; latency; virus replication;
KW Epstein-Barr virus; BHRF1; fusion protein; epitope tag;
KW drug screening; co-precipitation; ELISA; immunoassay; antibody;
KW protein interactive trapping; virucide; antitumour; diagnostic; ss.
OS Homo sapiens.
FH Key Location/Qualifiers
FT 5'UTR 1..200
FT /*tag= a
FT CDS 201..836
FT /*tag= b
FT /product= Bak protein
FT 3'UTR 837..2094
FT /*tag= c
PN WO9633416-A1.
PD 24-OCT-1996.
PF 19-APR-1996; U05639.
PR 20-APR-1995; US-426529.
PA (LXRB-) LXR BIOTECHNOLOGY INC.
PI Barr PJ, Kiefer MC;
DR WPI; 96-485886/48.
DR P-PSDB; W03668.
PT Screening for anti-viral agents - by detecting the ability of an
PT agent to disrupt the interaction of a Bak protein and a viral
PT protein
PS Disclosure; Fig 1; 24pp; English.
CC The sequence encodes Bak protein, which is a bcl-1 homologue which
CC interacts with Epstein-Barr virus (EBV) early lytic cycle BHRF1
CC protein, and is capable of modulating apoptosis. The Bak gene is
CC located on human chromosome-6 at 6p21-23. The protein may be used
CC in complete or partial form, or as an epitope tag fusion protein,
CC in a new virucide drug screening method, which involves combination
CC of Bak protein and a viral protein (e.g. EBV BHRF1), exposure to a
CC test compound, and monitoring for disruption of the interaction,
CC e.g. by co-precipitation, protein interactive trapping or ELISA.
CC Interaction of Bak and viral proteins allows viral replication or

BAD Dat

```
CC   latency in the absence of apoptosis.  Compounds which inhibit the
CC   interaction may be used as virucide, antitumour or diagnostic agents.
SQ   Sequence   2094 BP;    410 A;    608 C;    606 G;    470 T;

  Query Match          100.0%;  Score 2094;  DB 24;  Length 2094;
  Best Local Similarity 100.0%;  Pred. No. 0.00e+00;
  Matches  2094;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      1 gaggatctacaggggacaagtaaaggctacatccagatgccgggaatgcactgacgcccca 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      1 GAGGATCTACAGGGGACAAGTAAAGGCTACATCCAGATGCCGGGAATGCACTGACGCCCA 60

Db     61 ttcctggaaactgggctcccactcagcccctgggagcagcagccgccagcccctcggacc 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     61 TTCCTGGAAACTGGGCTCCCACTCAGCCCCTGGGAGCAGCAGCCGCCAGCCCCTCGGACC 120

Db    121 tccatctccaccctgctgagccacccgggttgggccaggatcccggcaggctgatcccgt 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    121 TCCATCTCCACCCTGCTGAGCCACCCGGGTTGGGCCAGGATCCCGGCAGGCTGATCCCGT 180

Db    181 cctccactgagacctgaaaaatggcttcggggcaaggcccaggtcctcccaggcaggagt 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    181 CCTCCACTGAGACCTGAAAAATGGCTTCGGGGCAAGGCCCAGGTCCTCCCAGGCAGGAGT 240

Db    241 gcggagagcctgccctgccctctgcttctgaggagcaggtagcccaggacacagaggagg 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    241 GCGGAGAGCCTGCCCTGCCCTCTGCTTCTGAGGAGCAGGTAGCCCAGGACACAGAGGAGG 300

Db    301 ttttccgcagctacgtttttttaccgccatcagcaggaacaggaggctgaaggggtggctg 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    301 TTTTCCGCAGCTACGTTTTTTTACCGCCATCAGCAGGAACAGGAGGCTGAAGGGGTGGCTG 360

Db    361 cccctgccgacccagagatggtcaccttacctctgcaacctagcagcaccatggggcagg 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    361 CCCCTGCCGACCCAGAGATGGTCACCTTACCTCTGCAACCTAGCAGCACCATGGGGCAGG 420

Db    421 tgggacggcagctcgccatcatcggggacgacatcaaccgacgctatgactcagagttcc 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    421 TGGGACGGCAGCTCGCCATCATCGGGGACGACATCAACCGACGCTATGACTCAGAGTTCC 480

Db    481 agaccatgttgcagcacctgcagcccacggcagagaatgcctatgagtacttcaccaaga 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    481 AGACCATGTTGCAGCACCTGCAGCCCACGGCAGAGAATGCCTATGAGTACTTCACCAAGA 540

Db    541 ttgccaccagcctgtttgagagtggcatcaattggggccgtgtggtggctcttctgggct 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    541 TTGCCACCAGCCTGTTTGAGAGTGGCATCAATTGGGGCCGTGTGGTGGCTCTTCTGGGCT 600

Db    601 tcggctaccgtctggccctacacgtctaccagcatggcctgactggcttcctaggccagg 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    601 TCGGCTACCGTCTGGCCCTACACGTCTACCAGCATGGCCTGACTGGCTTCCTAGGCCAGG 660

Db    661 tgacccgcttcgtggtcgacttcatgctgcatcactgcattgcccggtggattgcacaga 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    661 TGACCCGCTTCGTGGTCGACTTCATGCTGCATCACTGCATTGCCCGGTGGATTGCACAGA 720

Db    721 ggggtggctgggtggcagccctgaacttgggcaatggtcccatcctgaacgtgctggtgg 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    721 GGGGTGGCTGGGTGGCAGCCCTGAACTTGGGCAATGGTCCCATCCTGAACGTGCTGGTGG 780

Db    781 ttctgggtgtggttctgttgggccagtttgtggtacgaagattcttcaaatcatgactcc 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    781 TTCTGGGTGTGGTTCTGTTGGGCCAGTTTGTGGTACGAAGATTCTTCAAATCATGACTCC 840

Db    841 caagggtgccctttgggtcccggttcagacccctgcctggacttaagcgaagtctttgcc 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    841 CAAGGGTGCCCTTTGGGTCCCGGTTCAGACCCCTGCCTGGACTTAAGCGAAGTCTTTGCC 900

Db    901 ttctctgttcccttgcagggtccccccctcaagagtacagaagctttagcaagtgtgcact 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    901 TTCTCTGTTCCCTTGCAGGGTCCCCCCTCAAGAGTACAGAAGCTTTAGCAAGTGTGCACT 960

Db    961 ccagcttcggaggccctgcgtgggggccagtcaggctgcagaggcacctcaacattgcat 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    961 CCAGCTTCGGAGGCCCTGCGTGGGGGCCAGTCAGGCTGCAGAGGCACCTCAACATTGCAT 1020

Db   1021 ggtgctagtgccctctctctgggcccagggctgtggccgtctcctccctcagctctctgg 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1021 GGTGCTAGTGCCCTCTCTCTGGGCCCAGGGCTGTGGCCGTCTCCTCCCTCAGCTCTCTGG 1080

Db   1081 gacctccttagccctgtctgctaggcgctggggagactgataacttggggaggcaagaga 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1081 GACCTCCTTAGCCCTGTCTGCTAGGCGCTGGGGAGACTGATAACTTGGGGAGGCAAGAGA 1140

Db   1141 ctgggagccacttctccccagaaagtgtttaacggttttagctttttataataccccttgt 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1141 CTGGGAGCCACTTCTCCCCAGAAAGTGTTTAACGGTTTTAGCTTTTTATAATACCCTTGT 1200

Db   1201 gagagcccattcccaccattctacctgaggccaggacgtctggggtgtggggattggtgg 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1201 GAGAGCCCATTCCCACCATTCTACCTGAGGCCAGGACGTCTGGGGTGTGGGGATTGGTGG 1260

Db   1261 gtctatgttccccaggattcagctattctggaagatcagcaccctaagagatgggactag 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1261 GTCTATGTTCCCCAGGATTCAGCTATTCTGGAAGATCAGCACCCTAAGAGATGGGACTAG 1320

Db   1321 gacctgagcctggtcctggccgtccctaagcatgtgtcccaggagcaggacctactagga 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1321 GACCTGAGCCTGGTCCTGGCCGTCCCTAAGCATGTGTCCCAGGAGCAGGACCTACTAGGA 1380

Db   1381 gaggggggccaaggtcctgctcaactctacccctgctcccattcctccctccggccatac 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1381 GAGGGGGGCCAAGGTCCTGCTCAACTCTACCCCTGCTCCCATTCCTCCCTCCGGCCATAC 1440

Db   1441 tgcctttgcagttggactctcagggattctgggcttggggtgtggggtggggtggagtcg 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1441 TGCCTTTGCAGTTGGACTCTCAGGGATTCTGGGCTTGGGGTGTGGGGTGGGGTGGAGTCG 1500

Db   1501 cagaccagagctgtctgaactcacgtgtcagaagcctccaagcctgcctcccaaggtcct 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1501 CAGACCAGAGCTGTCTGAACTCACGTGTCAGAAGCCTCCAAGCCTGCCTCCCAAGGTCCT 1560

Db   1561 ctcagttctctcccttcctctctccttatagacacttgctcccaacccattcactacagg 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1561 CTCAGTTCTCTCCCTTCCTCTCTCCTTATAGACACTTGCTCCCAACCCATTCACTACAGG 1620

Db   1621 tgaaggctctcacccatccctgggggccttgggtgagtggcctgctaaggctcctccttg 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1621 TGAAGGCTCTCACCCATCCCTGGGGGCCTTGGGTGAGTGGCCTGCTAAGGCTCCTCCTTG 1680

Db   1681 cccagactacagggcttaggacttggtttgttatatcagggaaaaggagtagggagttca 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1681 CCCAGACTACAGGGCTTAGGACTTGGTTTGTTATATCAGGGAAAAGGAGTAGGGAGTTCA 1740

Db   1741 tctggagggttctaagtgggagaaggactatcaacaccactaggaatcccagaggtggat 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1741 TCTGGAGGGTTCTAAGTGGGAGAAGGACTATCAACACCACTAGGAATCCCAGAGGTGGAT 1800

Db   1801 cctccctcatggctctggcacagtgtaatccaggggtgtagatgggggaactgtgaatac 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1801 CCTCCCTCATGGCTCTGGCACAGTGTAATCCAGGGGTGTAGATGGGGGAACTGTGAATAC 1860

Db   1861 ttgaactctgttcccccacccctccatgctcctcacctgtctaggtctcctcagggtgggg 1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1861 TTGAACTCTGTTCCCCCACCCTCCATGCTCCTCACCTGTCTAGGTCTCCTCAGGGTGGGG 1920

Db   1921 ggtgacagtgccttctctattggcacagcctagggtcttgggggtcaggggggagaagtt 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1921 GGTGACAGTGCCTTCTCTATTGGCACAGCCTAGGGTCTTGGGGGTCAGGGGGGAGAAGTT 1980

Db   1981 cttgattcagccaaatgcagggagggggaggcagatggagcccataggccacccccctatcc 2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1981 CTTGATTCAGCCAAATGCAGGGAGGGGAGGCAGATGGAGCCCATAGGCCACCCCCTATCC 2040

Db   2041 tctgagtgtttggaaataaactgtgcaatcccctcaaaaaaaaaacggagatcc 2094
          |||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Qy   2041 TCTGAGTGTTTGGAAATAAACTGTGCAATCCCCTCAAAAAAAAAACGGAGATCC 2094


RESULT    2
ID   Q95492 standard; cDNA; 2072 BP.
AC   Q95492;
DT   21-NOV-1995  (first entry)
DE   Human Cdn-1 cDNA.
KW   Cdn-1; apoptosis modulator; adoptive immunotherapy; therapy; HIV;
KW   autoimmune disease; reperfusion injury; hepatitis, osteoporosis;
KW   shock; lymphoma; eczema; ss.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   CDS             188..903
FT   /*tag= a
PN   WO9515084-A.
PD   08-JUN-1995.
PF   30-NOV-1994; U13930.
PR   30-NOV-1993; US-160067.
PR   07-OCT-1994; US-320157.
PA   (LXRB-) LXR BIOTECHNOLOGY INC.
PI   Barr PJ,  Kiefer MC;
DR   WPI; 95-215106/28.
DR   P-PSDB; R77876.
PT   New nucleic acid sequences encoding Cdn apoptosis modulators  - and
PT   related vectors, transformed cells, proteins and antibodies, useful
PT   or diagnosis and treatment e.g. of HIV infection, reperfusion injury
PT   etc.
PS   Claim 4; Fig.3A-C; 66pp; English.
CC   Cdn-1 cDNA was isolated from a human heart cDNA library using a
CC   previously isolated clone as probe.  Recombinant Cdn-1 was produced
CC   in Sf9 and human colon adenocarcinoma HT29 cells.  Expression of
CC   Cdn-1 in WI-L2 lymphoblastoid cells resulted in increased cell
CC   survival in response to anti-Fas-mediated apoptosis.
SQ   Sequence   2072 BP;    404 A;    603 C;    598 G;    467 T;

  Query Match           98.9%;  Score 2072;  DB 14;  Length 2072;
  Best Local Similarity 100.0%;  Pred. No. 0.00e+00;
  Matches  2072;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      1 ggacaagtaaaggctacatccagatgccgggaatgcactgacgcccattcctggaaactg 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     14 GGACAAGTAAAGGCTACATCCAGATGCCGGGAATGCACTGACGCCCATTCCTGGAAACTG 73

Db     61 ggctcccactcagcccctgggagcagcagccgccagcccctcggacctccatctccaccc 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     74 GGCTCCCACTCAGCCCCTGGGAGCAGCAGCCGCCAGCCCCTCGGACCTCCATCTCCACCC 133

Db    121 tgctgagccacccgggttgggccaggatcccggcaggctgatcccgtcctccactgagac 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    134 TGCTGAGCCACCCGGGTTGGGCCAGGATCCCGGCAGGCTGATCCCGTCCTCCACTGAGAC 193

Db    181 ctgaaaaatggcttcggggcaaggcccaggtcctcccaggcaggagtgcggagagcctgc 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    194 CTGAAAAATGGCTTCGGGGCAAGGCCCAGGTCCTCCCAGGCAGGAGTGCGGAGAGCCTGC 253

Db    241 cctgccctctgcttctgaggagcaggtagcccaggacacagaggaggttttccgcagcta 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    254 CCTGCCCTCTGCTTCTGAGGAGCAGGTAGCCCAGGACACAGAGGAGGTTTTCCGCAGCTA 313

Db    301 cgtttttaccgccatcagcaggaacaggaggctgaaggggtggctgcccctgccgaccc 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    314 CGTTTTTTACCGCCATCAGCAGGAACAGGAGGCTGAAGGGGTGGCTGCCCCTGCCGACCC 373

Db    361 agagatggtcaccttacctctgcaacctagcagcaccatggggcaggtgggacggcagct 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    374 AGAGATGGTCACCTTACCTCTGCAACCTAGCAGCACCATGGGGCAGGTGGGACGGCAGCT 433

Db    421 cgccatcatcggggacgacatcaaccgacgctatgactcagagttccagaccatgttgca 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    434 CGCCATCATCGGGGACGACATCAACCGACGCTATGACTCAGAGTTCCAGACCATGTTGCA 493

Db    481 gcacctgcagcccacggcagagaatgcctatgagtacttcaccaagattgccaccagcct 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    494 GCACCTGCAGCCCACGGCAGAGAATGCCTATGAGTACTTCACCAAGATTGCCACCAGCCT 553

Db    541 gtttgagagtggcatcaattggggccgtgtggtggctcttctgggcttcggctaccgtct 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    554 GTTTGAGAGTGGCATCAATTGGGGCCGTGTGGTGGCTCTTCTGGGCTTCGGCTACCGTCT 613

Db    601 ggccctacacgtctaccagcatggcctgactggcttcctaggccaggtgacccgcttcgt 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    614 GGCCCTACACGTCTACCAGCATGGCCTGACTGGCTTCCTAGGCCAGGTGACCCGCTTCGT 673

Db    661 ggtcgacttcatgctgcatcactgcattgcccggtggattgcacagaggggtggctgggt 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    674 GGTCGACTTCATGCTGCATCACTGCATTGCCCGGTGGATTGCACAGAGGGGTGGCTGGGT 733

Db    721 ggcagccctgaacttgggcaatggtcccatcctgaacgtgctggtggttctgggtgtggt 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    734 GGCAGCCCTGAACTTGGGCAATGGTCCCATCCTGAACGTGCTGGTGGTTCTGGGTGTGGT 793

Db    781 tctgttgggccagtttgtggtacgaagattcttcaaatcatgactcccaagggtgccctt 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    794 TCTGTTGGGCCAGTTTGTGGTACGAAGATTCTTCAAATCATGACTCCCAAGGGTGCCCTT 853

Db    841 tgggtcccggttcagacccctgcctggacttaagcgaagtctttgccttctctgttccct 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    854 TGGGTCCCGGTTCAGACCCCTGCCTGGACTTAAGCGAAGTCTTTGCCTTCTCTGTTCCCT 913

Db    901 tgcagggtccccccctcaagagtacagaagctttagcaagtgtgcactccagcttcggagg 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    914 TGCAGGGTCCCCCCTCAAGAGTACAGAAGCTTTAGCAAGTGTGCACTCCAGCTTCGGAGG 973

Db    961 ccctgcgtgggggccagtcaggctgcagaggcacctcaacattgcatggtgctagtgccc 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    974 CCCTGCGTGGGGGCCAGTCAGGCTGCAGAGGCACCTCAACATTGCATGGTGCTAGTGCCC 1033

Db   1021 tctctctgggcccagggctgtggccgtctcctccctcagctctctgggacctccttagcc 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1034 TCTCTCTGGGCCCAGGGCTGTGGCCGTCTCCTCCCTCAGCTCTCTGGGACCTCCTTAGCC 1093

Db   1081 ctgtctgctaggcgctggggagactgataacttggggaggcaagagactgggagccactt 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1094 CTGTCTGCTAGGCGCTGGGGAGACTGATAACTTGGGGAGGCAAGAGACTGGGAGCCACTT 1153

Db   1141 ctccccagaaagtgtttaacggttttagctttttataataccccttgtgagagcccattcc 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1154 CTCCCCAGAAAGTGTTTAACGGTTTTAGCTTTTTATAATACCCTTGTGAGAGCCCATTCC 1213

Db   1201 caccattctacctgaggccaggacgtctggggtgtggggattggtgggtctatgttcccc 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1214 CACCATTCTACCTGAGGCCAGGACGTCTGGGGTGTGGGGATTGGTGGGTCTATGTTCCCC 1273

Db   1261 aggattcagctattctggaagatcagcaccctaagagatgggactaggacctgagcctgg 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1274 AGGATTCAGCTATTCTGGAAGATCAGCACCCTAAGAGATGGGACTAGGACCTGAGCCTGG 1333

Db   1321 tcctggccgtccctaagcatgtgtcccaggagcaggacctactaggagaggggggccaag 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1334 TCCTGGCCGTCCCTAAGCATGTGTCCCAGGAGCAGGACCTACTAGGAGAGGGGGGCCAAG 1393

Db   1381 gtcctgctcaactctacccctgctcccattcctccctccggccatactgcctttgcagtt 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1394 GTCCTGCTCAACTCTACCCCTGCTCCCATTCCTCCCTCCGGCCATACTGCCTTTGCAGTT 1453

Db   1441 ggactctcagggattctgggcttggggtgtggggtggggtggagtcgcagaccagagctg 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1454 GGACTCTCAGGGATTCTGGGCTTGGGGTGTGGGGTGGGGTGGAGTCGCAGACCAGAGCTG 1513

Db   1501 tctgaactcacgtgtcagaagcctccaagcctgcctcccaaggtcctctcagttctctcc 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1514 TCTGAACTCACGTGTCAGAAGCCTCCAAGCCTGCCTCCCAAGGTCCTCTCAGTTCTCTCC 1573

Db   1561 cttcctctctccttatagacacttgctcccaacccattcactacaggtgaaggctctcac 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Qy   1574 CTTCCTCTCTCCTTATAGACACTTGCTCCCAACCCATTCACTACAGGTGAAGGCTCTCAC 1633

Db   1621 ccatccctgggggccttgggtgagtggcctgctaaggctcctccttgcccagactacagg 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1634 CCATCCCTGGGGGCCTTGGGTGAGTGGCCTGCTAAGGCTCCTCCTTGCCCAGACTACAGG 1693

Db   1681 gcttaggacttggtttgttatatcagggaaaaggagtagggagttcatctggagggttct 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1694 GCTTAGGACTTGGTTTGTTATATCAGGGAAAAGGAGTAGGGAGTTCATCTGGAGGGTTCT 1753

Db   1741 aagtgggagaaggactatcaacaccactaggaatcccagaggtggatcctccctcatggc 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1754 AAGTGGGAGAAGGACTATCAACACCACTAGGAATCCCAGAGGTGGATCCTCCCTCATGGC 1813

Db   1801 tctggcacagtgtaatccaggggtgtagatgggggaactgtgaatacttgaactctgttc 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1814 TCTGGCACAGTGTAATCCAGGGGTGTAGATGGGGGAACTGTGAATACTTGAACTCTGTTC 1873

Db   1861 ccccaccctccatgctcctcacctgtctaggtctcctcagggtgggggtgacagtgcct 1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1874 CCCCACCCTCCATGCTCCTCACCTGTCTAGGTCTCCTCAGGGTGGGGGGTGACAGTGCCT 1933

Db   1921 tctctattggcacagcctagggtcttgggggtcaggggggagaagttcttgattcagcca 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1934 TCTCTATTGGCACAGCCTAGGGTCTTGGGGGTCAGGGGGGAGAAGTTCTTGATTCAGCCA 1993

Db   1981 aatgcagggaggggaggcagatggagcccataggccacccccctatcctctgagtgtttgg 2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1994 AATGCAGGGAGGGGAGGCAGATGGAGCCCATAGGCCACCCCCTATCCTCTGAGTGTTTGG 2053

Db   2041 aaataaactgtgcaatcccctcaaaaaaaaaa 2072
          |||||||||||||||||||||||||||||||
Qy   2054 AAATAAACTGTGCAATCCCCTCAAAAAAAAAA 2085


RESULT    3
ID    Q95493 standard; DNA; 6511 BP.
AC    Q95493;
DT    21-NOV-1995  (first entry)
DE    Human Cdn-2 DNA.
KW    Cdn-2; apoptosis modulator; adoptive immunotherapy; therapy; HIV;
KW    autoimmune disease; reperfusion injury; hepatitis, osteoporosis;
KW    shock; lymphoma; eczema; ss.
OS    Homo sapiens.
FH    Key             Location/Qualifiers
FT    CDS             3312..3947
FT    /*tag= a
PN    WO9515084-A.
PD    08-JUN-1995.
PF    30-NOV-1994; U13930.
PR    30-NOV-1993; US-160067.
PR    07-OCT-1994; US-320157.
PA    (LXRB-) LXR BIOTECHNOLOGY INC.
PI    Barr PJ,  Kiefer MC;
DR    WPI; 95-215106/28.
DR    P-PSDB; R77877.
PT    New nucleic acid sequences encoding Cdn apoptosis modulators  - and
PT    related vectors, transformed cells, proteins and antibodies, useful
PT    or diagnosis and treatment e.g. of HIV infection, reperfusion injury
PT    etc.
PS    Claim 6; Fig.5A-H; 66pp; English.
CC    Cdn-2 cDNA was isolated from a human placental genomic library
CC    using a 950 bp fragment of Cdn-1 cDNA.  Expression of Cdn-2
CC    in mouse progenitor B-cell FL5.12 cells decreased IL-3-induced
CC    apoptosis.  The Cdn-2 protein displayed 97% amino acid identity
CC    with Cdn-1 (R77876).
SQ    Sequence   6511 BP;    1513 A;    1620 C;    1605 G;    1773 T;

  Query Match           90.7%;  Score 1899;  DB 14;  Length 6511;
  Best Local Similarity  97.3%;  Pred. No. 0.00e+00;
  Matches  2026;  Conservative    0;  Mismatches  43;  Indels  14;  Gaps  11;

Db   3123 gggacaagtaaaggctacatccagatgctgggaatgcactgacgcccattcctggaaact 3182
          ||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Qy     13 GGGACAAGTAAAGGCTACATCCAGATGCCGGGAATGCACTGACGCCCATTCCTGGAAACT 72

Db   3183 gggctcccactcagcccctgggagcagcagccgccagcccctcgggacctccatctccac 3242
          |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Qy     73 GGGCTCCCACTCAGCCCCTGGGAGCAGCAGCCGCCAGCCCCTCGG-ACCTCCATCTCCAC 131

Db   3243 cctgctgagccacccgggttgggccaggatcccggcaggctgatcccgtcctccactgag 3302
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    132 CCTGCTGAGCCACCCGGGTTGGGCCAGGATCCCGGCAGGCTGATCCCGTCCTCCACTGAG 191

Db   3303 acctgaaaaatggcttcggggcaaggcccaggtcctcccaggcaggagtgcggagagcct 3362
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    192 ACCTGAAAAATGGCTTCGGGGCAAGGCCCAGGTCCTCCCAGGCAGGAGTGCGGAGAGCCT 251

Db   3363 gccctgccctctgcttctgaggagcaggtagcccaggacacagaggaggttttccgcagc 3422
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    252 GCCCTGCCCTCTGCTTCTGAGGAGCAGGTAGCCCAGGACACAGAGGAGGTTTTCCGCAGC 311

Db   3423 tacgttttttaccaccatcagcaggaacaggaggctgaaggggcggctgcccctgccgac 3482
          |||||||||||| |||||||||||||||||||||||||||| ||||||||||||||||
Qy    312 TACGTTTTTTACCGCCATCAGCAGGAACAGGAGGCTGAAGGGGTGGCTGCCCCTGCCGAC 371

Db   3483 ccagagatggtcaccttacctctgcaacctagcagcaccatggggcaggtgggacggcag 3542
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    372 CCAGAGATGGTCACCTTACCTCTGCAACCTAGCAGCACCATGGGGCAGGTGGGACGGCAG 431

Db   3543 ctcgccatcattggggacgacatcaaccgacgctatgactcagagttccagaccatgttg 3602
          |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Qy    432 CTCGCCATCATCGGGGACGACATCAACCGACGCTATGACTCAGAGTTCCAGACCATGTTG 491

Db   3603 cagcacctgcagcccacggcagagaatgcctatgagtacttcaccaagattgcctccagc 3662
          |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Qy    492 CAGCACCTGCAGCCCACGGCAGAGAATGCCTATGAGTACTTCACCAAGATTGCCACCAGC 551

Db   3663 ctgtttgagagtggcatcaattggggccgtgtggtggctcttctgggcttcagctaccgt 3722
          |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Qy    552 CTGTTTGAGAGTGGCATCAATTGGGGCCGTGTGGTGGCTCTTCTGGGCTTCGGCTACCGT 611

Db   3723 ctggccctacacatctaccagcgtggcctgactggcttcctgggccaggtgacccgcttt 3782
          |||||||||| |||||||| |||||||||||||||||| |||||||||||||||||||
Qy    612 CTGGCCCTACACGTCTACCAGCATGGCCTGACTGGCTTCCTAGGCCAGGTGACCCGCTTC 671

Db   3783 gtggtggacttcatgctgcatcactgcattgcccggtggattgcacagaggggtggctgg 3842
          ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    672 GTGGTCGACTTCATGCTGCATCACTGCATTGCCCGGTGGATTGCACAGAGGGGTGGCTGG 731

Db   3843 gtggcagccctgaacttgggcaatggtcccatcctgaacgtgctggtggttctgggtgtg 3902
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    732 GTGGCAGCCCTGAACTTGGGCAATGGTCCCATCCTGAACGTGCTGGTGGTTCTGGGTGTG 791

Db   3903 gttctgttgggccagtttgtggtacgaagattcttcaaatcatgactcccaagggtgccc 3962
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    792 GTTCTGTTGGGCCAGTTTGTGGTACGAAGATTCTTCAAATCATGACTCCCAAGGGTGCCC 851

Db   3963 tttggggtcccagttcagacccctgcctggacttaagcgaagtctttgccttctctgctc 4022
          |||||| |||| ||||||||||||||||||||||||||||||||||||||||||| ||
Qy    852 TTTGGG-TCCCGGTTCAGACCCCTGCCTGGACTTAAGCGAAGTCTTTGCCTTCTCTGTTC 910

Db   4023 cttgcaggggtccccccctcaagagtacagaagctttagcaagtgtgcactccagcttcgg 4082
          | |    ||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    911 CCTTGCAGGGTCCCCCCTCAAGAGTACAGAAGCTTTAGCAAGTGTGCACTCCAGCTTCGG 970

Db   4083 agggcccctgtgtgggggccagtcaggctgcagaggcacctcaacattccatggtgctag 4142
          |||  ||||| |||||||||||||||||||||||||||||||||||| ||||||||||
Qy    971 AGG--CCCTGCGTGGGGGCCAGTCAGGCTGCAGAGGCACCTCAACATTGCATGGTGCTAG 1028

Db   4143 tgggccctctctctgggcccagggggctgtggcgtctcctccctcagctctctgggacctc 4202
          ||  |||||||||||||||||||     | ||||||||||||||||||||||||||||||
Qy   1029 TG--CCCTCTCTCTGGGCCCAGGGCTGTGGCCGTCTCCTCCCTCAGCTCTCTGGGACCTC 1086

Db   4203 cttagccctgtctgctaggcgctggggagactgataacttggggaggcaagagactggga 4262
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Qy   1087 CTTAGCCCTGTCTGCTAGGCGCTGGGGAGACTGATAACTTGGGGAGGCAAGAGACTGGGA 1146

Db   4263 gccacttctccccagaaagtgtttaatggttttagctttttataatacccttgtgagagc 4322
          |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Qy   1147 GCCACTTCTCCCCAGAAAGTGTTTAACGGTTTTAGCTTTTTATAATACCCTTGTGAGAGC 1206

Db   4323 ccattcccaccattctacctgaggccaggacgtctggggtgtggggattggtgtgtctat 4382
          |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Qy   1207 CCATTCCCACCATTCTACCTGAGGCCAGGACGTCTGGGGTGTGGGGATTGGTGGGTCTAT 1266

Db   4383 gttccccaggattcagctattctggaagatcagcaccctaagagatgggactaggacctg 4442
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1267 GTTCCCCAGGATTCAGCTATTCTGGAAGATCAGCACCCTAAGAGATGGGACTAGGACCTG 1326

Db   4443 agcctggtcctggccgtccctaagcatgtgtcccaggagcaggacctactaggagagggg 4502
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1327 AGCCTGGTCCTGGCCGTCCCTAAGCATGTGTCCCAGGAGCAGGACCTACTAGGAGAGGGG 1386

Db   4503 ggccaaggtcctgctcaactctacccctgctcccattcctccctccggccatactgcctt 4562
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1387 GGCCAAGGTCCTGCTCAACTCTACCCCTGCTCCCATTCCTCCCTCCGGCCATACTGCCTT 1446

Db   4563 tgcagttggactctcagggattctgggcttggggtgtggggtggggtggagtcg-agacc 4621
          |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Qy   1447 TGCAGTTGGACTCTCAGGGATTCTGGGCTTGGGGTGTGGGGTGGGGTGGAGTCGCAGACC 1506

Db   4622 agagctgtctgaactcatgtgtcagaagccctccaagcctgcctcccagggtcctctcag 4681
          |||||||||||||||| |||||||||||| |||||||||||||||| ||||||||||||
Qy   1507 AGAGCTGTCTGAACTCACGTGTCAGAAGCC-TCCAAGCCTGCCTCCCAAGGTCCTCTCAG 1565

Db   4682 ttctctcccttcctctctccttatagacacttgctcccaacccattcactacaggtgaag 4741
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1566 TTCTCTCCCTTCCTCTCTCCTTATAGACACTTGCTCCCAACCCATTCACTACAGGTGAAG 1625

Db   4742 gctcctcaccccccatccctgggccttgggtgagtaacctgctaaggc--ctccttgccca 4799
          |||| ||||||    |   ||||||||||||||  ||||||||||  ||||||||||||
Qy   1626 GCTC-TCACCCATCCCTGGGGGCCTTGGGTGAGTGGCCTGCTAAGGCTCCTCCTTGCCCA 1684

Db   4800 gactacagggcttaggacttggtttgttatttcagggaaaaggagtagggagttcatctg 4859
          |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Qy   1685 GACTACAGGGCTTAGGACTTGGTTTGTTATATCAGGGAAAAGGAGTAGGGAGTTCATCTG 1744

Db   4860 gagggttctaagtgggagaaggactatcaacaccactaggaatcccagaggtgggatcct 4919
          |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Qy   1745 GAGGGTTCTAAGTGGGAGAAGGACTATCAACACCACTAGGAATCCCAGAGGTGG-ATCCT 1803

Db   4920 ccctcatggctctggcacagtgtaatccaggggtgtagatgggggaactgtgaatacttg 4979
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1804 CCCTCATGGCTCTGGCACAGTGTAATCCAGGGGTGTAGATGGGGGAACTGTGAATACTTG 1863

Db   4980 aactctgttcccccacctccatgctcctcacctgtctaggtctcctcagggtggggggt 5039
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1864 AACTCTGTTCCCCCACCCTCCATGCTCCTCACCTGTCTAGGTCTCCTCAGGGTGGGGGGT 1923

Db   5040 gagagtgccttctctattgggcacagcctagggtcttgggggtc-gggggagaagttct 5098
          || |||||||||||||||| |||||||||||||||||||||| ||||||||||||||||
Qy   1924 GACAGTGCCTTCTCTATTGG-CACAGCCTAGGGTCTTGGGGGTCAGGGGGGAGAAGTTCT 1982

Db   5099 tgattcagccaaatgcagggaggggaggcagatggagcccataggccacctcctatcctc 5158
          |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Qy   1983 TGATTCAGCCAAATGCAGGGAGGGGAGGCAGATGGAGCCCATAGGCCACCCCCTATCCTC 2042

Db   5159 tgagtgtttggaaataaactgtgcaatcccctcaaaaaaataa 5201
          |||||||||||||||||||||||||||||||||||| ||
Qy   2043 TGAGTGTTTGGAAATAAACTGTGCAATCCCCTCAAAAAAAAAA 2085


RESULT    4
ID    T17375 standard; cDNA; 1968 BP.
AC    T17375;
DT    02-JUL-1996  (first entry)
DE    Bcl-Y cDNA.
KW    Bcl-Y; apoptosis; cell proliferation; cell death; diagnosis;
KW    gene therapy; ss.
OS    Homo sapiens.
FH    Key             Location/Qualifiers
FT    CDS             65..700
FT    /*tag=  a
PN    WO9605232-A1.
PD    22-FEB-1996.
PF    09-AUG-1995; U10103.
PR    09-AUG-1994; US-287427.
PR    11-OCT-1994; US-321071.
PA    (IMMU-) IMMUNOGEN INC.
PI    Chittenden TD;
DR    WPI; 96-139648/14.
DR    P-PSDB; R81451.
PT    New isolated human Bcl-Y protein - used to develop prods. for
PT    treating disorders characterised by inappropriate cell proliferation
PT    or cell death
PS    Claim 6; Fig 4; 100pp; English.
CC    A full-length cDNA clone (T17375) codes for Bcl-Y (R81451), a
CC    protein that induces apoptosis in cells and functions as a negative
CC    regulator of Bcl-2 function.  It was isolated from a Jurkat cell
CC    cDNA library using as probe a partial Bcl-Y cDNA clone obtd. by
CC    PCR of DNA derived from the Namalwa cell line.  The cDNA can be used
CC    for prodn. of recombinant Bcl-Y, as a probe, to produce transgenic
CC    animal models, and in the gene therapy of disorders characterised by
CC    inappropriate cell proliferation or cell death.
SQ    Sequence   1968 BP;    382 A;    560 C;    576 G;    450 T;

  Query Match           90.1%;  Score 1886;  DB 19;  Length 1968;
  Best Local Similarity  99.5%;  Pred. No. 0.00e+00;
  Matches  1940;  Conservative    0;  Mismatches   0;  Indels   9;  Gaps   7;

Db      1 tgagccacccgggttgggccaggatcccggcaggctgatcccgtcctccactgagacctg 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    137 TGAGCCACCCGGGTTGGGCCAGGATCCCGGCAGGCTGATCCCGTCCTCCACTGAGACCTG 196

Db     61 aaaaatggcttcggggcaaggcccaggtcctcccaggcaggagtgcggagagcctgccct 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    197 AAAAATGGCTTCGGGGCAAGGCCCAGGTCCTCCCAGGCAGGAGTGCGGAGAGCCTGCCCT 256

Db    121 gccctctgcttctgaggagcaggtagcccaggacacagaggaggttttccgcagctacgt 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    257 GCCCTCTGCTTCTGAGGAGCAGGTAGCCCAGGACACAGAGGAGGTTTTCCGCAGCTACGT 316

Db    181 tttttaccgccatcagcaggaacaggaggctgaaggggtggctgcccctgccgacccaga 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    317 TTTTTACCGCCATCAGCAGGAACAGGAGGCTGAAGGGGTGGCTGCCCCTGCCGACCCAGA 376

Db    241 gatggtcaccttacctctgcaacctagcagcaccatggggcaggtgggacggcagctcgc 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    377 GATGGTCACCTTACCTCTGCAACCTAGCAGCACCATGGGGCAGGTGGGACGGCAGCTCGC 436

Db    301 catcatcggggacgacatcaaccgacgctatgactcagagttccagaccatgttgcagca 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    437 CATCATCGGGGACGACATCAACCGACGCTATGACTCAGAGTTCCAGACCATGTTGCAGCA 496

Db    361 cctgcagcccacggcagagaatgcctatgagtacttcaccaagattgccaccagcctgtt 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    497 CCTGCAGCCCACGGCAGAGAATGCCTATGAGTACTTCACCAAGATTGCCACCAGCCTGTT 556

Db    421 tgagagtggcatcaattggggccgtgtggtggctcttctgggcttcggctaccgtctggc 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    557 TGAGAGTGGCATCAATTGGGGCCGTGTGGTGGCTCTTCTGGGCTTCGGCTACCGTCTGGC 616

Db    481 cctacacgtctaccagcatggcctgactggcttcctaggccaggtgacccgcttcgtggt 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    617 CCTACACGTCTACCAGCATGGCCTGACTGGCTTCCTAGGCCAGGTGACCCGCTTCGTGGT 676

Db    541 cgacttcatgctgcatcactgcattgcccggtggattgcacagaggggtggctgggtggc 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    677 CGACTTCATGCTGCATCACTGCATTGCCCGGTGGATTGCACAGAGGGGTGGCTGGGTGGC 736

Db    601 agccctgaacttgggcaatggtcccatcctgaacgtgctggtggttctgggtgtggttct 660
```

```
Qy    737 AGCCCTGAACTTGGGCAATGGTCCCATCCTGAACGTGCTGGTGGTTCTGGGTGTGGTTCT 796

Db    661 gttgggccagtttgtggtacgaagattcttcaaatcatgactcccaagggtgccctttgg 720
Qy    797 GTTGGGCCAGTTTGTGGTACGAAGATTCTTCAAATCATGACTCCCAAGGGTGCCCTTTGG 856

Db    721 gtcccggttcagacccctgcctggacttaagcgaagtctttgccttctctgttcccttgc 780
Qy    857 GTCCCGGTTCAGACCCCTGCCTGGACTTAAGCGAAGTCTTTGCCTTCTCTGTTCCCTTGC 916

Db    781 aggggtccccccctcaagagtacagaagctttagcaagtgtgcactccagcttcggagggc 840
Qy    917 AGGG-TCCCCCCTCAAGAGTACAGAAGCTTTAGCAAGTGTGCACTCCAGCTTCGGAGG-- 973

Db    841 ccctgcgtgggggccagtcaggctgcagaggcacctcaacattgcatggtgctagtgggc 900
Qy    974 CCCTGCGTGGGGGCCAGTCAGGCTGCAGAGGCACCTCAACATTGCATGGTGCTAGTG--C 1031

Db    901 cctctctctgggcccaggggctgtggccgtctcctccctcagctctctgggacctcctta 960
Qy   1032 CCTCTCTCTGGGCCCAGGG-CTGTGGCCGTCTCCTCCCTCAGCTCTCTGGGACCTCCTTA 1090

Db    961 gccctgtctgctaggcgctggggagactgataacttggggaggcaagagactgggagcca 1020
Qy   1091 GCCCTGTCTGCTAGGCGCTGGGGAGACTGATAACTTGGGGAGGCAAGAGACTGGGAGCCA 1150

Db   1021 cttctccccagaaagtgtttaacggttttagctttttataatacccttgtgagagcccat 1080
Qy   1151 CTTCTCCCCAGAAAGTGTTTAACGGTTTTAGCTTTTTATAATACCCTTGTGAGAGCCCAT 1210

Db   1081 tcccaccattctacctgaggccaggacgtctggggtgtggggattggtgggtctatgttc 1140
Qy   1211 TCCCACCATTCTACCTGAGGCCAGGACGTCTGGGGTGTGGGGATTGGTGGGTCTATGTTC 1270

Db   1141 cccaggattcagctattctggaagatcagcaccctaagagatgggactaggacctgagcc 1200
Qy   1271 CCCAGGATTCAGCTATTCTGGAAGATCAGCACCCTAAGAGATGGGACTAGGACCTGAGCC 1330

Db   1201 tggtcctggccgtccctaagcatgtgtcccaggagcaggacctactaggagaggggggcc 1260
Qy   1331 TGGTCCTGGCCGTCCCTAAGCATGTGTCCCAGGAGCAGGACCTACTAGGAGAGGGGGGCC 1390

Db   1261 aaggtcctgctcaactctacccctgctcccattcctccctccggccatactgcctttgca 1320
Qy   1391 AAGGTCCTGCTCAACTCTACCCCTGCTCCCATTCCTCCCTCCGGCCATACTGCCTTTGCA 1450

Db   1321 gttggactctcagggattctgggcttggggtgtggggtggggtggagtcgcagaccagag 1380
Qy   1451 GTTGGACTCTCAGGGATTCTGGGCTTGGGGTGTGGGGTGGGGTGGAGTCGCAGACCAGAG 1510

Db   1381 ctgtctgaactcacgtgtcagaagcctccaagcctgcctcccaaggtcctctcagttctc 1440
Qy   1511 CTGTCTGAACTCACGTGTCAGAAGCCTCCAAGCCTGCCTCCCAAGGTCCTCTCAGTTCTC 1570

Db   1441 tcccttcctctctccttatagacacttgctcccaacccattcactacaggtgaaggctct 1500
Qy   1571 TCCCTTCCTCTCTCCTTATAGACACTTGCTCCCAACCCATTCACTACAGGTGAAGGCTCT 1630

Db   1501 caccccatccctgggggccttgggtgagtggcctgctaaggctcctccttgcccagacta 1560
Qy   1631 CACCC-ATCCCTGGGGGCCTTGGGTGAGTGGCCTGCTAAGGCTCCTCCTTGCCCAGACTA 1689

Db   1561 cagggcttaggacttggtttgttatatcagggaaaaggagtagggagttcatctggaggg 1620
Qy   1690 CAGGGCTTAGGACTTGGTTTGTTATATCAGGGAAAAGGAGTAGGGAGTTCATCTGGAGGG 1749

Db   1621 ttctaagtgggagaaggactatcaacaccactaggaatcccagaggtgggatcctccctc 1680
Qy   1750 TTCTAAGTGGGAGAAGGACTATCAACACCACTAGGAATCCCAGAGGTGG-ATCCTCCCTC 1808

Db   1681 atggctctggcacagtgtaatccaggggtgtagatgggggaactgtgaatacttgaactc 1740
Qy   1809 ATGGCTCTGGCACAGTGTAATCCAGGGGTGTAGATGGGGGAACTGTGAATACTTGAACTC 1868

Db   1741 tgttcccccaccctccatgctcctcacctgtctaggtctcctcagggtgggggggtgacag 1800
Qy   1869 TGTTCCCCCACCCTCCATGCTCCTCACCTGTCTAGGTCTCCTCAGGGTGGGGGGTGACAG 1928

Db   1801 tgccttctctattgggcacagcctagggtcttgggggtcaggggggagaagttcttgatt 1860
Qy   1929 TGCCTTCTCTATTGG-CACAGCCTAGGGTCTTGGGGGTCAGGGGGGAGAAGTTCTTGATT 1987

Db   1861 cagccaaatgcagggaggggaggcagatggagcccataggccacccccctatcctctgagt 1920
Qy   1988 CAGCCAAATGCAGGGAGGGGAGGCAGATGGAGCCCATAGGCCACCCCCTATCCTCTGAGT 2047

Db   1921 gtttggaaataaactgtgcaatcccctca 1949
Qy   2048 GTTTGGAAATAAACTGTGCAATCCCCTCA 2076
```

RESULT 5
ID Q95494 standard; DNA; 5408 BP.
AC Q95494;
DT 21-NOV-1995 (first entry)
DE Human Cdn-3 DNA.
KW Cdn-3; apoptosis modulator; adoptive immunotherapy; therapy; HIV;
KW autoimmune disease; reperfusion injury; hepatitis, osteoporosis;
KW shock; lymphoma; eczema; ss.
OS Homo sapiens.
FH Key Location/Qualifiers
FT CDS 1665..1931
FT /*tag= a
PN WO9515084-A.
PD 08-JUN-1995.
PF 30-NOV-1994; U13930.
PR 30-NOV-1993; US-160067.
PR 07-OCT-1994; US-320157.
PA (LXRB-) LXR BIOTECHNOLOGY INC.
PI Barr PJ, Kiefer MC;
DR WPI; 95-215106/28.
DR P-PSDB; R77878.
PT New nucleic acid sequences encoding Cdn apoptosis modulators - and
PT related vectors, transformed cells, proteins and antibodies, useful
PT or diagnosis and treatment e.g. of HIV infection, reperfusion injury
PT etc.
PS Disclosure; Fig.7A-G; 66pp; English.
CC Southern blot analysis of human genomic DNA and a panel of
CC human/rodent somatic cell DNAs revealed at least 3 Cdn-related
CC genes residing on chromosomes 6, 11 and 20. Cdn-3 (R77878)
CC did not contain the structural features of Cdn-1 (R77876), Cdn-2
CC (R77877) or other Bcl-2 family members.
SQ Sequence 5408 BP; 1369 A; 1384 C; 1314 G; 1341 T;

Query Match 71.4%; Score 1496; DB 14; Length 5408;
Best Local Similarity 89.6%; Pred. No. 0.00e+00;
Matches 1880; Conservative 0; Mismatches 186; Indels 33; Gaps 23;

```
Db   1476 gggacaagcaaaggctacatccagataccagaaatgcactgacgcccgttcctggaagct 1535
Qy     13 GGGACAAGTAAAGGCTACATCCAGATGCCGGGAATGCACTGACGCCCATTCCTGGAAACT 72

Db   1536 gggctcccactcagcccctgggagcagcagcctccagccccttgggaccttcaactccac 1595
Qy     73 GGGCTCCCACTCAGCCCCTGGGAGCAGCAGCCGCCAGCCCCTCGG-ACCTCCATCTCCAC 131

Db   1596 cctgctgacccacgcgggttgagccagcatccctggaggctgacactgtcctccactgag 1655
Qy    132 CCTGCTGAGCCACCCGGGTTGGGCCAGGATCCCGGCAGGCTGATCCCGTCCTCCACTGAG 191

Db   1656 acctgaaaaaatggcatcggggcaaggcccagggcctcccaggcaggagtgcggaaagcct 1715
Qy    192 ACCTGAAAAAATGGCTTCGGGGCAAGGCCCAGGTCCTCCCAGGCAGGAGTGCGGAGAGCCT 251

Db   1716 gccctgccctctgcttctgaggagcaggtagcccaggacatggagg-ggttttccgcagc 1774
```

```
Qy    252 GCCCTGCCCTCTGCTTCTGAGGAGCAGGTAGCCCAGGACACAGAGGAGGTTTTCCGCAGC 311

Db   1775 tacgtttttttaccaccatcagcaggaacaggaggctgaaggggcggccgcccctgccgac 1834

Qy    312 TACGTTTTTTACCGCCATCAGCAGGAACAGGAGGCTGAAGGGGTGGCTGCCCCTGCCGAC 371

Db   1835 ccagagatggtcaccttgcccctccaacctagcagcaccatggggcaggtgggacggcag 1894

Qy    372 CCAGAGATGGTCACCTTACCTCTGCAACCTAGCAGCACCATGGGGCAGGTGGGACGGCAG 431

Db   1895 ctcgccatcaccagg-acgacatcaaccggcactatgacttcggagttccagaccatgct 1953

Qy    432 CTCGCCATCATCGGGGACGACATCAACCGACGCTATGACT-CAGAGTTCCAGACCATGTT 490

Db   1954 gcagcacctgcagcccacggcagagaagcgcctacgagtacttcaccaagatcgcctccag 2013

Qy    491 GCAGCACCTGCAGCCCACGGCAGAGAATGCCTATGAGTACTTCACCAAGATTGCCACCAG 550

Db   2014 cctgtttgagagtggcatcaaccggggccgtgtggtggctctcctgggcttcggctaccg 2073

Qy    551 CCTGTTTGAGAGTGGCATCAATTGGGGCCGTGTGGTGGCTCTTCTGGGCTTCGGCTACCG 610

Db   2074 tctggtcctacatgtctaccagcacggcttgactggcttcctgggcctggtgacccgctt 2133

Qy    611 TCTGGCCCTACACGTCTACCAGCATGGCCTGACTGGCTTCCTAGGCCAGGTGACCCGCTT 670

Db   2134 cgtggtc---ttcatgctgcaacaaggcatcgcccggtggatctcgcagaggggcggctg 2190

Qy    671 CGTGGTCGACTTCATGCTGCATCACTGCATTGCCCGGTGGATTGCACAGAGGGGTGGCTG 730

Db   2191 ggtggcagccctggacttgggcaatagtcccatcctgaacgtgctggtggttgtgggtgt 2250

Qy    731 GGTGGCAGCCCTGAACTTGGGCAATGGTCCCATCCTGAACGTGCTGGTGGTTCTGGGTGT 790

Db   2251 ggttctgctgggccagtttgtggtaagaagattcttcaaatcatgactcccaggggtgtc 2310

Qy    791 GGTTCTGTTGGGCCAGTTTGTGGTACGAAGATTCTTCAAATCATGACTCCCAAGGGTGCC 850

Db   2311 ctttggggtcccagctgtgaccccctgcctggacttaagccaagtctttgccttccccact 2370

Qy    851 CTTTGGG-TCCCGGTTCAGACCCCTGCCTGGACTTAAGCGAAGTCTTTGCCTTCTCTGTT 909

Db   2371 cccttgcaggggtcacccttcaaaagtacagaagctctagcaagtgtgcacccccgctgc 2430

Qy    910 CCCTTGCAGGG-TCCCCCCTCAAGAGTACAGAAGCTTTAGCAAGTGTGCACTCCAGCTTC 968

Db   2431 ggagggcccctgcgtgggggccagtcaggctgcggaggcacctcaacattgcacggtgct 2490

Qy    969 GGAGG--CCCTGCGTGGGGGCCAGTCAGGCTGCAGAGGCACCTCAACATTGCATGGTGCT 1026

Db   2491 agtgggccctctctctgggcccaggggctgtgccctcctcccttggctctctgggacctc 2550

Qy   1027 AGTGCCCTCTCTCTGGGCCCAGGGCTGTGGCCGTCTCCTCCCTCAGCTCTCTGGGACCTC 1086

Db   2551 cttagtcttgtctgctaggcgctgcagaggctgataacttggggaagcaagagactggga 2610

Qy   1087 CTTAGCCCTGTCTGCTAGGCGCTGGGGAGACTGATAACTTGGGGAGGCAAGAGACTGGGA 1146

Db   2611 gccactcctccccagtaagtgtttaacggttttagctttttataataccccttgggagagc 2670

Qy   1147 GCCACTTCTCCCCAGAAAGTGTTTAACGGTTTTAGCTTTTTATAATACCCTTGTGAGAGC 1206

Db   2671 ccattcccaccattctacccaaggccgggatgtctggggtgtgggggttggtgggtcgta 2730

Qy   1207 CCATTCCCACCATTCTACCTGAGGCCAGGACGTCTGGGGTGTGGGGATTGGTGGGTC-TA 1265

Db   2731 acctacgtgccccaggattcagctattctggaagatcag-agcctaagagctaggacttg 2789

Qy   1266 ---T--GTTCCCCAGGATTCAGCTATTCTGGAAGATCAGCACCCTAAGAGATGGGACTAG 1320

Db   2790 atcctggtcctgg-cc-gtccctaagcat-catgtgtcccaggagcaggactgactggga 2846
```

```
Qy   1321 GACCTGAGCCTGGTCCTGGCCGTCCCTAAGCATGTGTCCCAGGAGCAGGACCTACTAGGA 1380

Db   2847 gaggggga-ccaaggtcctacccagctctccccgtgccccccattcctcc-tccggccatac 2904

Qy   1381 GAGGGGGGCCAAGGTCCTGCTCAACTCTACCCCTGCTCCCATTCCTCCCTCCGGCCATAC 1440

Db   2905 tgcctttgcagttggactctcagggattctgggcttggggtgtggggcggcgtggagtaa 2964

Qy   1441 TGCCTTTGCAGTTGGACTCTCAGGGATTCTGGGCTTGGGGTGTGGGGTGGGGTGGAGTCG 1500

Db   2965 caggccagagctgtctgaacttatgtgtcagaagcctccaagcctgcctcccaaggtcct 3024

Qy   1501 CAGACCAGAGCTGTCTGAACTCACGTGTCAGAAGCCTCCAAGCCTGCCTCCCAAGGTCCT 1560

Db   3025 ctcagctctctcccttcctctctccttatagatacttgctcccaacccattcactacagg 3084

Qy   1561 CTCAGTTCTCTCCCTTCCTCTCTCCTTATAGACACTTGCTCCCAACCCATTCACTACAGG 1620

Db   3085 tgaaggccctcacccatccctgggggccttgggtgagtgatgcgctaaggcccctccccg 3144

Qy   1621 TGAAGGCTCTCACCCATCCCTGGGGGCCTTGGGTGAGTGGCCTGCTAAGGCTCCTCCTTG 1680

Db   3145 cccagactacagggcttggtttagggcttggtttgttatttcagggataaggagtaggga 3204

Qy   1681 CCCAGACTACAGGGCTT-----AGGACTTGGTTTGTTATATCAGGGAAAAGGAGTAGGGA 1735

Db   3205 gttcatctggaaggttctaagtgggagaaggactatcaacaccac-aggaatcccagagg 3263

Qy   1736 GTTCATCTGGAGGGTTCTAAGTGGGAGAAGGACTATCAACACCACTAGGAATCCCAGAGG 1795

Db   3264 tgggatcctccctcatggctctggcacagtgtaatccaggggtggagataggg-aactgt 3322

Qy   1796 TGG-ATCCTCCCTCATGGCTCTGGCACAGTGTAATCCAGGGGTGTAGATGGGGGAACTGT 1854

Db   3323 gaatacctgaactctgtcccccgaccctccatgctcctcacctttctgggtctctcctca 3382

Qy   1855 GAATACTTGAACTCTGTTCCCCCACCCTCCATGCTCCTCACCTGTCTAGGTCTCCTCAGG 1914

Db   3383 gtgtgggggtgagagtaccttctctatcgggcacagcctagggtgttgggggtgaagggg 3442

Qy   1915 GTG-GGGGGTGACAGTGCCTTCTCTATTGG-CACAGCCTAGGGTCTTGGGGGTCAGGGGG 1972

Db   3443 gagaagttcttgattcagccaaatgcagggagggggaggcagaaggagccccacaggccact 3502

Qy   1973 GAGAAGTTCTTGATTCAGCCAAATGCAGGGAGGGGAGGCAGATGGAGCCCATAGGCCACC 2032

Db   3503 ccctatcctctgagtgtttggaaataaactgtgcaatcccatcaaaaaaaaaaaggaga 3561

Qy   2033 CCCTATCCTCTGAGTGTTTGGAAATAAACTGTGCAATCCCCTCAAAAAAAAAACGGAGA 2091
```

```
RESULT    6
ID   T42139 standard; cDNA; 1286 BP.
AC   T42139;
DT   22-FEB-1997  (first entry)
DE   Bak-2 gene.
KW   Human; Bak-2; apoptosis; latency; virus replication;
KW   Epstein-Barr virus; BHRF1; fusion protein; epitope tag;
KW   drug screening; co-precipitation; ELISA; immunoassay; antibody;
KW   protein interactive trapping; virucide; antitumour; diagnostic; ds.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   5'UTR           1..543
FT   /*tag=  a
FT   CDS             544..1179
FT   /*tag=  b
FT   /product= Bak-2 protein
FT   3'UTR           1180..1286
FT   /*tag=  c
PN   WO9633416-A1.
PD   24-OCT-1996.
PF   19-APR-1996; U05639.
PR   20-APR-1995; US-426529.
```

```
PA   (LXRB-) LXR BIOTECHNOLOGY INC.
PI   Barr PJ,  Kiefer MC;
DR   WPI; 96-485886/48.
DR   P-PSDB; W03669.
PT   Screening for anti-viral agents - by detecting the ability of an
PT   agent to disrupt the interaction of a Bak protein and a viral
PT   protein
PS   Disclosure; Fig 2; 24pp; English.
CC   The sequence encodes Bak-2 protein, which is a bcl-1 homologue which
CC   interacts with Epstein-Barr virus (EBV) early lytic cycle BHRF1
CC   protein, and is capable of modulating apoptosis.  The Bak-2 gene is
CC   located on human chromosome-20.  The protein may be used in
CC   complete or partial form, or as an epitope tag fusion protein, in a
CC   new virucide drug screening method, which involves combination of
CC   Bak-2 protein and a viral protein (e.g. EBV BHRF1), exposure to a
CC   test compound, and monitoring for disruption of the interaction,
CC   e.g. by co-precipitation, protein interactive trapping or ELISA.
CC   Interaction of Bak-2 and viral proteins allows viral replication or
CC   latency in the absence of apoptosis.  Compounds which inhibit the
CC   interaction may be used as virucide, antitumour or diagnostic agents.
SQ   Sequence   1286 BP;    298 A;    341 C;    328 G;    319 T;

  Query Match           42.2%;  Score 883;  DB 24;  Length 1286;
  Best Local Similarity  98.1%;  Pred. No. 0.00e+00;
  Matches   916;  Conservative    0;  Mismatches  15;  Indels   3;  Gaps   3;

Db    355 gggacaagtaaaggctacatccagatgctgggaatgcactgacgcccattcctggaaact 414
          ||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Qy     13 GGGACAAGTAAAGGCTACATCCAGATGCCGGGAATGCACTGACGCCCATTCCTGGAAACT 72

Db    415 gggctcccactcagcccctgggagcagcagccgccagcccctcgggacctccatctccac 474
          |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Qy     73 GGGCTCCCACTCAGCCCCTGGGAGCAGCAGCCGCCAGCCCCTCGG-ACCTCCATCTCCAC 131

Db    475 cctgctgagccacccgggttgggccaggatcccggcaggctgatcccgtcctccactgag 534
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    132 CCTGCTGAGCCACCCGGGTTGGGCCAGGATCCCGGCAGGCTGATCCCGTCCTCCACTGAG 191

Db    535 acctgaaaaatggcttcggggcaaggcccaggtcctcccaggcaggagtgcggagagcct 594
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    192 ACCTGAAAAATGGCTTCGGGGCAAGGCCCAGGTCCTCCCAGGCAGGAGTGCGGAGAGCCT 251

Db    595 gccctgccctctgcttctgaggagcaggtagcccaggacacagaggaggttttccgcagc 654
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    252 GCCCTGCCCTCTGCTTCTGAGGAGCAGGTAGCCCAGGACACAGAGGAGGTTTTCCGCAGC 311

Db    655 tacgtttttttaccaccatcagcaggaacaggaggctgaaggggcggctgcccctgccgac 714
          |||||||||||||| |||||||||||||||||||||||||||| ||||||||||||||||
Qy    312 TACGTTTTTTACCGCCATCAGCAGGAACAGGAGGCTGAAGGGGTGGCTGCCCCTGCCGAC 371

Db    715 ccagagatggtcaccttacctctgcaacctagcagcaccatggggcaggtgggacggcag 774
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    372 CCAGAGATGGTCACCTTACCTCTGCAACCTAGCAGCACCATGGGGCAGGTGGGACGGCAG 431

Db    775 ctcgccatcattggggacgacatcaaccgacgctatgactcagagttccagaccatgttg 834
          |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Qy    432 CTCGCCATCATCGGGGACGACATCAACCGACGCTATGACTCAGAGTTCCAGACCATGTTG 491

Db    835 cagcacctgcagcccacggcagagaatgcctatgagtacttcaccaagattgcctccagc 894
          |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Qy    492 CAGCACCTGCAGCCCACGGCAGAGAATGCCTATGAGTACTTCACCAAGATTGCCACCAGC 551

Db    895 ctgtttgagagtggcatcaattggggccgtgtggtggctcttctgggcttcagctaccgt 954
          |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Qy    552 CTGTTTGAGAGTGGCATCAATTGGGGCCGTGTGGTGGCTCTTCTGGGCTTCGGCTACCGT 611

Db    955 ctggccctacacatctaccagcgtggcctgactggcttcctgggccaggtgacccgcttt 1014
          |||||||||||| ||||||||| |||||||||||||||||| ||||||||||||||||||
Qy    612 CTGGCCCTACACGTCTACCAGCATGGCCTGACTGGCTTCCTAGGCCAGGTGACCCGCTTC 671

Db   1015 gtggtggacttcatgctgcatcactgcattgcccggtggattgcacagaggggtggctgg 1074
          ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    672 GTGGTCGACTTCATGCTGCATCACTGCATTGCCCGGTGGATTGCACAGAGGGGTGGCTGG 731

Db   1075 gtggcagccctgaacttgggcaatggtcccatcctgaacgtgctggtggttctgggtgtg 1134
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    732 GTGGCAGCCCTGAACTTGGGCAATGGTCCCATCCTGAACGTGCTGGTGGTTCTGGGTGTG 791

Db   1135 gttctgttgggccagtttgtggtacgaagattcttcaaatcatgactcccaagggtgcct 1194
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    792 GTTCTGTTGGGCCAGTTTGTGGTACGAAGATTCTTCAAATCATGACTCCCAAGGGTGCCC 851

Db   1195 ttggggtcccagttcagacccctgcctggacttaagcgaagtctttgccttctctgctcc 1254
          || ||||||| |||||||||||||||||||||||||||||||||||||||||||||| |||
Qy    852 TTTGGGTCCCGGTTCAGACCCCTGCCTGGACTTAAGCGAAGTCTTTGCCTTCTCTGTTCC 911

Db   1255 -ttgcagg-tccccccctcaagagtacagaagctt 1286
           ||||||| ||||||||||||||||||||||||||
Qy    912 CTTGCAGGGTCCCCCCTCAAGAGTACAGAAGCTT 945


RESULT    7
ID   Q10572 standard; DNA; 1047 BP.
AC   Q10572;
DT   09-APR-1991  (first entry)
DE   Human Natriuretic Peptide Receptor B.
KW   NPRB; ANP; BNP; CNP; kidney failure; heart failure; protein kinase;
KW   hyperaldosteronism; glaucoma; guanyl cyclase.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   Peptide         1..22
FT   /label= signal sequence
FT   Protein         12
FT   /label= mature NPBR
FT   Domain          23..455
FT   /label= extracellular domain
FT   /note= "binds natriuretic peptides A,B and C]"
FT   Domain          456..456
FT   /label= transmembrane domain
FT   Domain          479..1047
FT   /label= cytoplasmic domain
FT   /note= "GC and protien kinase activity"
FT   Modified -site  24..26
FT   /label= N-glycos_site
FT   Modified -site  35..37
FT   /label= N-glycos_site
FT   Modified -site  161..163
FT   /label= N-glycos_site
FT   Modified -site  195..197
FT   /label= N-glycos_site
FT   Modified -site  244..246
FT   /label= N-glycos_site
FT   Modified -site  277..279
FT   /label= N-glycos_site
FT   Modified -site  349..351
FT   /label= N-glycos_site
FT   Modified -site  600..602
FT   /label= N-glycos_site
PN   WO9100292-A.
PD   10-JAN-1991.
PF   22-JUN-1990; U03586.
PR   23-JUN-1989; US-370673.
PA   (GETH ) GENENTECH INC.
PI   Chang M, Goeddel D, Lowe D;
DR   WPI; 91-036711/05.
DR   N-PSDB; Q10324.
PT   Natriuretic protein receptor B - for diagnosis and treatment of
PT   kidney failure, heart failure, hyperaldosteronism, glaucoma etc.
PS   Claim 3; Fig 1; 49pp; English.
CC   The sequence was derived from the DNA encoding natriuretic peptide
CC   receptor B, NPRB, having guanyl cyclase (GC) activity and protein
CC   kinase activity.  The DNA can be inserted into expression vectors
CC   for the prodn. of the protein, opt. after being mutated to produce
CC   NPRB analogues. The protein has a mol wt. of 115 kD (calculated Mr=
CC   114,952).   The protein (or variants) can be used in treatment of
CC   natriuretic peptide disorders, and also to isolate peptides using
```

```
CC   affinity chromatography.  Antibodies with affinity for NPRB can
CC   also be prepd.
SQ   Sequence   1047 BP;    87 A;    15 C;    83 G;    51 T;

  Query Match              4.5%;  Score 95;  DB 2;  Length 1047;
  Best Local Similarity    9.1%;  Pred. No. 4.06e-46;
  Matches    93;  Conservative  293;  Mismatches 618;  Indels  13;  Gaps  13;

Db      17 vrnngarnntnavvnnnhnnsyawawnrvgnavanavnangrannvdnrnvssnnngacs 76
             :: ||:     ::       :: : : ::|   |  : |  :|   : : :::    ||
Cp    1225 GGTAGAATGGTGGGAATGGGCTCTCACAAGGGTATTATAAAAAGCTAAAACCGTTAAACA 1166

Db      77 nynannsavdnknyhdndnnngn-gcvynaasvarnashwrnnnnt-agavasgnsaknd 134
              :  : :: :    : :        | ::    ::     ::::    | || : :  :|: :
Cp    1165 CTTTCTGGGGAGAAGTGGCTCCCAGTCTCTTGCCTCCCCAAGTTATCAGTCTCCCCAGCG 1106

Db     135 hyrtnvrtgnsankngnnvvtnhghnnwtaraannyndartddrnhyntnngvnnanngs 194
             ::    :: :| :| : |    :: :|      |:|    : :|: ::: :         : |   :
Cp    1105 CCTAGCA-GACAGG-GCTAAGGAGGTCCCAGAGAGCTGAGGGAGGAGACGGCCACAGCCC 1048

Db     195 nnsvnhnvyarnnggnnnathnnrangrnvync-gnnnmnhnnnnnanrnnntngdyvnn 253
               :: : : |:    |    |  :    |  : ::    |       :              :      : :
Cp    1047 TGGGCCCAGAGAGAGGGCACTAGCACCATGCAATGTTGAGGTGCCTCTGCAGCCTGACTG 988

Db     254 yndvngnsnragntratgrnwndnrtrnnananrnannntvnvntyrnnnnnnynnnnnrn 313
              :    :  ||    |   : :  |:      : |  |:       :          :
Cp     987 GCCCCCACGCAGGGCCTCCGAAGC-TGGAGTGCACACTTGCTAAAGCTTCTGTACTCTTG 929

Db     314 nnrarndngvnngnsnmnnnagcnydgnnnyanvnnntnnggtrndgnrnvnkmngrry 373
               : : :   :   |  : :    | |  :|         :          |  |   :|    : :: |::
Cp     928 AGGGGGGACCCTG-CAAGGGAACAGAGAAGGCAAAGACTTCGCTTAAGTCCAGGCAGGGG 870

Db     374 hgvtgnvvmdkndrntdnvnwamgdndsgdnnnaahysganknnwwtgrnnnwvkgann 433
             : :|| ::::        :      :     : |:            ::: |   :  ::     :   ::: |
Cp     869 T-CTGAACCGGGACCCAAAGGGCACCCTTGGGAGTCATG-ATTTGAAGAATCTTCGTACC 812

Db     434 sdnnncandnddnscdktnnstnanvangtgntnnmngvssnnnnrknmnnknnasmnwr 493
                   |   :        |::    :   |  :   |       : |: :      :: :  :    :: ::
Cp     811 ACAAACTGGCCCAACAGAACCACACCCAGAACCACCAGCACGTTCAGGATGGGACCATTG 752

Db     494 nrwnnnnngnsnryhkgagsrntnsnrgssygsnmtahgkynnnantghnkgnvvankhv 553
                       : ::::    :       : : ::: :   |  :   :          |       ::| : :
Cp     751 CCCAAGTTCAGGGCTGCCACCCAGCCACCCCTCTGTGCAATCCACCGGGCAATGCAGTGA 692

Db     554 nkkrnnntrnvnnnnkhmrdvnnnhntrnngacndnnnncnvtnycnrgsnndnnnndsn 613
              :  :     |: :        ::: :     :    :       | :          :     | :|      :       ::
Cp     691 TGCAGCATGAAGTCGACCACGAAGCGGGTCACCTGGCCTAGGAAGCCAGTCAGGCCATGC 632

Db     614 nndwmnrysnnndnvkgmannhnsnnsshgsnkssncvvdsrnvnkntdygnasnrstan 673
              ::: :::     : :: :|   : :   :::|   :::   :::::   :      ::    :   : |
Cp     631 TGGTAGACGTGTAGGGCCAGACGGTAGCCGAAGCCCAGAAGAGCCACCACACGGCCCCAA 572

Db     674 nddnnanyakknntannnnsgnnnnttgmnaadvysngnnnnnnanrsgnnynngndnsn 733
             ::         :|    :     |        :|   | |: ||:::    |              |        :
Cp     571 TTGATGCCACTCTCAAACAGGCTGGTGGC-AATCTTGGTGAAGTACTCATAGGCATTCTC 513

Db     734 knnvnkvrngnrnynrnsndrtnnnnnnvnnmnrcwandnanrndngnnkgnnrrnnkng 793
             :   : :::     :     : : ::                   :         :                   |    :: :
Cp     512 TGCCGTGGGCTGCAGGTGCTGCAACATGGTCTGGAACTCTGAGTCATAGCGTCGGTTGAT 453

Db     794 gtsnndnnnnrmnnyannnnkn-vnnrtnaynnnkrkanannynnnnhsvannnkrgntv 852
             ||:           :      :           : :  :|    :     ::                 :::        : |   :
Cp     452 GTCGTCCCCGATGATGGCGAGCTGCCGTCCCACCTGCCCCATGGTGCTGCTAGGTTGCAG 393

Db     853 nanandsvtnynsdnvgntansanstnmnvvtnnndnytcndanndnndvykvntngday 912
                   :::|     :: :    |    :    :    :  ::       :  : |            :    ::::: |   : :
Cp     392 AGGTAAGGTGACCATCTCTGGGTCGGCAGGG-GCAGCCACCC-CTTCAGCCTCCTGTTCC 335

Db     913 mvvsgnngrngnrhannarmananndavssnrnrhrnhdnnrnrngvhtgnvcagvvgn 972
             :: |   |: |   :|    |              : :: :   :   ::   :   :: |  :  |   :  |
Cp     334 TGCTGATGGCGGTAAAAAACGTAGCTGCGGAAAACCTCCTCTGTGTCCTGGGCTACCTGC 275

Db     973 kmnrycnngdtvntasrmnsngnanknhvssttkdandnngcnnnnnrgdvnmkgkg 1029
             ::  :     : :    ::: :  |  |  : ::  :       :      ||          :|::  ::|:|
Cp     274 TCCTCAGAAGCAGAGGGCAGGGCAGGCTCTCCGCACTCCTGCCTGGGAGGACCTGGG 218

RESULT    8
ID   Q10572 standard; DNA; 1047 BP.
AC   Q10572;
DT   09-APR-1991  (first entry)
DE   Human Natriuretic Peptide Receptor B.
KW   NPRB; ANP; BNP; CNP; kidney failure; heart failure; protein kinase;
KW   hyperaldosteronism; glaucoma; guanyl cyclase.
OS   Homo sapiens.
FH   Key              Location/Qualifiers
FT   Peptide          1..22
FT   /label= signal sequence
FT   Protein          12
FT   /label= mature NPBR
FT   Domain           23..455
FT   /label= extracellular domain
FT   /note= "binds natriuretic peptides A,B and C]"
FT   Domain           456..456
FT   /label= transmembrane domain
FT   Domain           479..1047
FT   /label= cytoplasmic domain
FT   /note= "GC and protien kinase activity"
FT   Modified -site  24..26
FT   /label= N-glycos_site
FT   Modified -site  35..37
FT   /label= N-glycos_site
FT   Modified -site  161..163
FT   /label= N-glycos_site
FT   Modified -site  195..197
FT   /label= N-glycos_site
FT   Modified -site  244..246
FT   /label= N-glycos_site
FT   Modified -site  277..279
FT   /label= N-glycos_site
FT   Modified -site  349..351
FT   /label= N-glycos_site
FT   Modified -site  600..602
FT   /label= N-glycos_site
PN   WO9100292-A.
PD   10-JAN-1991.
PF   22-JUN-1990; U03586.
PR   23-JUN-1989; US-370673.
PA   (GETH ) GENENTECH INC.
PI   Chang M, Goeddel D, Lowe D;
DR   WPI; 91-036711/05.
DR   N-PSDB; Q10324.
PT   Natriuretic protein receptor B - for diagnosis and treatment of
PT   kidney failure, heart failure, hyperaldosteronism, glaucoma etc.
PS   Claim 3; Fig 1; 49pp; English.
CC   The sequence was derived from the DNA encoding natriuretic peptide
CC   receptor B, NPRB, having guanyl cyclase (GC) activity and protein
CC   kinase activity.  The DNA can be inserted into expression vectors
CC   for the prodn. of the protein, opt. after being mutated to produce
CC   NPRB analogues. The protein has a mol wt. of 115 kD (calculated Mr=
CC   114,952).  The protein (or variants) can be used in treatment of
CC   natriuretic peptide disorders, and also to isolate peptides using
CC   affinity chromatography.  Antibodies with affinity for NPRB can
CC   also be prepd.
SQ   Sequence   1047 BP;    87 A;    15 C;    83 G;    51 T;

  Query Match              3.7%;  Score 78;  DB 2;  Length 1047;
  Best Local Similarity    8.8%;  Pred. No. 2.35e-34;
  Matches    72;  Conservative  234;  Mismatches 503;  Indels  10;  Gaps  10;

Db     194 snnsvnhnvyarnnggnnnathnnrangrnvyncgnnnmnhnnnnnanrnnntngdyvnn 253
             : :: : ::      |        :  :  |: :: ||    : :          :    |  ::
Qy     209 GGGGCAAGGCCCAGGTCCTCCCAGGCAGGAGTGCGGAGAGCCTGCCCTGCCCTCTGCTTC 268

Db     254 yndvngnsnragntratgrnwndnrtrnnananrnanntvnvnt-yrnnnnnnynnnnnr 312
             : ::   : :           : : : : :   |           : : | ::                   :
```

```
Qy    269 TGAGGAGCAGGTAGCCCAGGACACAGAGGAGGTTTTCCGCAGCTACGTTTTTTACCGCCA 328

Db    313 nnnrarndngvnngnsnmnnnagcnydgnnnyanvnnntnnnggtrndgnrnvnkmngrr 372

Qy    329 TCAGCAGGAACAGGAGGCTGAAGGGGTGGCTGCCCCTGCCGACCCAGAG-ATGGTCACCT 387

Db    373 yhgvtgnvvmdknndrntdnvnwamgdndsgdnnnaahysgankнnwwtgrnnnwvkgan 432

Qy    388 TACCTCTGCAACCTAGCAGCACCATGGGGCAGGTGGGACGGCAGCTCGCCATCATCGGGG 447

Db    433 nsdnnncandnddnscdktnnstnanvangtgntnnmngvssnnnnrknmnnknnasmnw 492

Qy    448 ACGACATCAACCGAC-GCTATGACTCAGAGTTCCAGACCATGTTGCAGCACCTGCAGCCC 506

Db    493 rnrwnnnnngnsnryhkgagsrntnsnrgssygsnmtahgkynnnantghnkgnvvankh 552

Qy    507 ACG-GCAGAGAATGCCTATGAGTACTTCACCAAGATTGCCACCAGCCTGTTTGAGAGTGG 565

Db    553 vnkkrnnntrnvnnnnkhmrdvnnnhntrnngacndnnnncnvtnycnrgsnndnnnnd- 611

Qy    566 CATCAATTGGGGCCGTGTGGTGGCTCTTCTGGGCTTCGGCTACCGTCTGGCCCTACACGT 625

Db    612 snnndwmnrysnnndnvkgmannhnsnnsshgsnkssncvvdsrnvnkntdygnasnrst 671

Qy    626 CTACCAGCATGGCCTGACTGGCTTCCTAGGCCAGGTGACCCGCTTCGTGGTCG-ACTTCA 684

Db    672 annddnnanyakknntannnnsgnnnnttgmnaadvysngnnnnnnnanrsgnnynngndn 731

Qy    685 TGCTGCATCACTGCATTGCCCGGTGGATTGCACAGAGGGGTGGCTGGGTGGCAGCCCTGA 744

Db    732 snknnvnkvrngnrnynrnsndrtnnnnnnvnnmnrcwandnanrndngnnkgnnrrnnk 791

Qy    745 ACTTGGGCAATGGTCCCATCCTGAACGTGCTGGTGGTTCTGGGTGTGGTTCTGTTGGGCC 804

Db    792 nggt-snndnnnnrmnnyannnnknvnnrtnaynnnkrkanannynnnnhsvannnkrgn 850

Qy    805 AGTTTGTGGTACGAAGATTCTTCAAATCATGACTCCCAAGGGTGCCCTTTGGGTCCCGGT 864

Db    851 tvnanandsvtnynsdnvgntansanstnmnvvtnnndnytcndanndnndvy-kvntng 909

Qy    865 TC-AGACCCCTGCCTGGACTTAAGCGAAGTCTTTGCCTTCTCTGTTCCCTTGCAGGGTCC 923

Db    910 daymvvsgnngrngnrhannnarmananndavssnrnrhrnhdnnrnrngvhtgnvcagv 969

Qy    924 CCCCTCAAGAGTACAGAAGCTTTAGCAAGTGTGCACTCCAGCTTCGGAGGCCCTGCGTGG 983

Db    970 vgnkmnrycnngdtvnt-asrmnsngnanknhvssttkd 1007

Qy    984 GGGCCAGTCAGGCTGCAGAGGCACCTCAACATTGCATGG 1022
```

```
RESULT    9
ID    Q51746 standard; cDNA; 91 BP.
AC    Q51746;
DT    31-MAY-1994  (first entry)
DE    Oligonucleotide probe MK14-A
KW    Oligonucleotide; DNA probe; mycobacteria; disease diagnosis;
KW    ss.
OS    Synthetic.
PN    EP-571911-A.
PD    01-DEC-1993.
PF    24-MAY-1993; 108325.
PR    26-MAY-1992; US-889651.
PA    (BECT ) BECTON DICKINSON CO.
PI    Shank DD,  Spears PA;
DR    WPI; 93-378844/48.
PT    New oligo:nucleotide probes specific for Mycobacteria - used for
PT    detection and amplification of Mycobacteria nucleic acid in
PT    samples
PS    Claim 3; Page 14; 23pp; English.
CC    Oligonucleotide probe MK14-A consists of nucleotides 5-95 of MK14
CC    (Q51735).  It hybridized to all spp. of mycobacteria tested, but
CC    cross reacted to a few non-mycobacterial spp.  The probe may
CC    be useful as an initial screen for mycobacterial infection.
CC    See also Q51735-45 and Q51747-59.
SQ    Sequence   91 BP;    5 A;    17 C;    15 G;    4 T;

  Query Match            2.1%;  Score 43;  DB 9;  Length 91;
  Best Local Similarity  10.2%;  Pred. No. 1.89e-11;
  Matches     6;  Conservative   45;  Mismatches   8;  Indels   0;  Gaps   0;

Db      3 ctccggcgssvhsyyvvhvvshhhsvhhvvhhvhvsvvvvhhvvhvvhhvhyhvyvsvc 61

Qy    120 CTCCATCTCCACCCTGCTGAGCCACCCGGGTTGGGCCAGGATCCCGGCAGGCTGATCCC 178


RESULT   10
ID    N81164 standard; DNA; 204 BP.
AC    N81164;
DT    08-NOV-1990  (first entry)
DE    Base substituted E.coli beta-galactosidase alpha-fragment.
KW    E.coli beta galactosidase alpha-fragment; base substitutions; ss.
OS    Escherichia coli.
FH    Key             Location/Qualifiers
FT    misc_feature    19..69
FT    /*tag=  a
FT    /function=multiple cloning site
FT    primer_bind     187..204
FT    /*tag=  b
PN    EP-285123-A.
PD    05-MAY-1988.
PF    30-MAR-1988; 105163.
PR    03-APR-1987; US-034819.
PA    (SUSO) SUOMEN SOKERI OY.
PI    Lehtovaara P, Knowles J, Koivula A, Bamford J, Reinikainen T;
DR    WPI; 88-279927/40.
PT    Introducing random point mutations into nucleic acods -
PT    by prepn of single stranded template, annealing a primer, elongation,
PT    misincorporation, completion of molecules and screening.
PS    Disclosure; p; English.
CC    Random point mutations were introduced into the alpha fragment of
CC    E.coli beta-galactosidase. The wild type sequence was obtained as a
CC    single stranded template and an oligonucleotide was hybridised to
CC    it to generate a popn of DNA molecules which terminate at all
CC    possible nucleotide positions within a specified region. The
CC    variable 3' ends generated in this way are used as primers for
CC    reverse transcriptase. Nucleotides are misincorporated by the
CC    transcriptase and the molecules are completed to forms that can be
CC    amplified and then expressed in a suitable host-vector system.
CC    The sequence covers all 176 difft base substitutions, most of which
CC    occurred singularly in any given mutant.
CC    See also P80575.
SQ    Sequence   204 BP;   21 A;   47 C;   17 G;   11 T;   108 Others;

  Query Match            2.1%;  Score 44;  DB 1;  Length 204;
  Best Local Similarity  11.2%;  Pred. No. 4.60e-12;
  Matches    13;  Conservative   60;  Mismatches  43;  Indels   0;  Gaps   0;

Db     77 cdchvgccgymrttthhyrrmrbnvyrdynrsdaaawyccyrrsvkydccynachhddhy 136

Qy   1520 CTCACGTGTCAGAAGCCTCCAAGCCTGCCTCCCAAGGTCCTCTCAGTTCTCTCCCTTCCT 1579

Db    137 vybbbvynvhnhnncncccbnnhvchnvhbnnhrnwayvrhdarrddvhccvchcc 192

Qy   1580 CTCTCCTTATAGACACTTGCTCCCAACCCATTCACTACAGGTGAAGGCTCTCACCC 1635


RESULT   11
ID    N81164 standard; DNA; 204 BP.
AC    N81164;
DT    08-NOV-1990  (first entry)
DE    Base substituted E.coli beta-galactosidase alpha-fragment.
KW    E.coli beta galactosidase alpha-fragment; base substitutions; ss.
OS    Escherichia coli.
FH    Key             Location/Qualifiers
FT    misc_feature    19..69
```

```
FT   /*tag= a
FT   /function=multiple cloning site
FT   primer_bind     187..204
FT   /*tag= b
PN   EP-285123-A.
PD   05-MAY-1988.
PF   30-MAR-1988; 105163.
PR   03-APR-1987; US-034819.
PA   (SUSO) SUOMEN SOKERI OY.
PI   Lehtovaara P, Knowles J, Koivula A, Bamford J, Reinikainen T;
DR   WPI; 88-279927/40.
PT   Introducing random point mutations into nucleic acods -
PT   by prepn of single stranded template, annealing a primer, elongation,
PT   misincorporation, completion of molecules and screening.
PS   Disclosure; p; English.
CC   Random point mutations were introduced into the alpha fragment of
CC   E.coli beta-galactosidase. The wild type sequence was obtained as a
CC   single stranded template and an oligonucleotide was hybridised to
CC   it to generate a popn of DNA molecules which terminate at all
CC   possible nucleotide positions within a specified region. The
CC   variable 3' ends generated in this way are used as primers for
CC   reverse transcriptase. Nucleotides are misincorporated by the
CC   transcriptase and the molecules are completed to forms that can be
CC   amplified and then expressed in a suitable host-vector system.
CC   The sequence covers all 176 difft base substitutions, most of which
CC   occurred singularly in any given mutant.
CC   See also P80575.
SQ   Sequence   204 BP;   21 A;   47 C;   17 G;   11 T;   108 Others;

  Query Match            2.1%;  Score 45;  DB 1;  Length 204;
  Best Local Similarity  10.4%;  Pred. No. 1.11e-12;
  Matches    10;  Conservative   56;  Mismatches  29;  Indels   1;  Gaps   1;

Db     95 rrmrbnvyrdynrsdaaawyccyrrsvkydccynachhddhyvybbbvynvhnhnncncc 154
          ::: : : :::  :: | ::  :::: :::  :  |:::: :::::::: :: :  | ||
Cp   1544 GGCTTGGAGGCTTCTGACACG-TGAGTTCAGACAGCTCTGGTCTGCGACTCCACCCCACC 1486

Db    155 cbnnhvchnvhbnnhrnwayvrhdarrddvhccvch 190
          |:  ::|: :::  :: :|:: :: :::: :| :|:
Cp   1485 CCACACCCCAAGCCCAGAATCCCTGAGAGTCCAACT 1450


RESULT   12
ID   Q51746 standard; cDNA; 91 BP.
AC   Q51746;
DT   31-MAY-1994  (first entry)
DE   Oligonucleotide probe MK14-A
KW   Oligonucleotide; DNA probe; mycobacteria; disease diagnosis;
KW   ss.
OS   Synthetic.
PN   EP-571911-A.
PD   01-DEC-1993.
PF   24-MAY-1993; 108325.
PR   26-MAY-1992; US-889651.
PA   (BECT ) BECTON DICKINSON CO.
PI   Shank DD,  Spears PA;
DR   WPI; 93-378844/48.
PT   New oligo:nucleotide probes specific for Mycobacteria - used for
PT   detection and amplification of Mycobacteria nucleic acid in
PT   samples
PS   Claim 3; Page 14; 23pp; English.
CC   Oligonucleotide probe MK14-A consists of nucleotides 5-95 of MK14
CC   (Q51735).  It hybridized to all spp. of mycobacteria tested, but
CC   cross reacted to a few non-mycobacterial spp.  The probe may
CC   be useful as an initial screen for mycobacterial infection.
CC   See also Q51735-45 and Q51747-59.
SQ   Sequence   91 BP;    5 A;    17 C;    15 G;    4 T;

  Query Match            2.0%;  Score 42;  DB 9;  Length 91;
  Best Local Similarity  5.8%;  Pred. No. 7.69e-11;
  Matches     3;  Conservative   44;  Mismatches   5;  Indels   0;  Gaps   0;

Db     12 svhsyyvvhvvshhhsvhhvvhhvhvsvvvvhhvvhvvhhvhyhvyvsvctc 63
          :::::::::::: ::: :::::::: ::::::::::::::: ::::::: :|||
Cp   1778 GGTGTTGATAGTCCTTCTCCCACTTAGAACCCTCCAGATGAACTCCCTACTC 1727


RESULT   13
ID   Q70465 standard; DNA; 114 BP.
AC   Q70465;
DT   05-APR-1995  (first entry)
DE   Generic DNA sequence to generate a random TSAR petide library.
KW   TSAR; totally synthetic affinity reagent; synthetic; binding domain;
KW   effector domain; concateneated heterofunctional protein; linker;
KW   direct; rapid; detection; screening; treatment; generic; ss.
OS   Synthetic.
FH   Key             Location/Qualifiers
FT   misc_feature    55..60
FT   /*tag= a
FT   /note= "this sequence represents 'Z'; Z can be a
FT   sequence of 6, 9 or 12 nucleotides (see
FT   comments)"
PN   WO9418318-A.
PD   18-AUG-1994.
PF   01-FEB-1994; U00977.
PR   01-FEB-1993; US-013416.
PR   30-DEC-1993; US-176500.
PR   31-JAN-1994; US-189331.
PA   (UYNC-) UNIV NORTH CAROLINA.
PI   Fowlkes DM,  Kay BK;
DR   WPI; 94-279739/34.
DR   P-PSDB; R65150 and R65151.
PT   Identifying proteins or peptide(s) which bind a ligand - by
PT   screening a recombinant vector library expressing fusion proteins
PT   comprising a binding domain and an effector domain
PS   Disclosure; Page 35; 255pp; English.
CC   Q70465 is a generic DNA sequence used to generate random TSAR (Totally
CC   Synthetic Affinity Reagents) peptides.This generic formula can also be
CC   represented  as follows: X(NNB)6(TGC)(NNB)11Z(NNB)14(TGC)(NNB)3Y. X
CC   and Y are flanking  restriction sites (X is not the same as Y) that are
CC   not specified further. Other generic sequences are shown in Q70466-68.
CC   Other specific peptides generated by these generic sequences are shown in
CC   R65151-54. TSARs are concatenated heterofunctional proteins or peptides,
CC   comprising at least two functional regions - a binding domain with
CC   affinity for a ligand and a second effector peptide portion that is
CC   chemically or biologically active.They may further comprise a linker
CC   peptide between the 2 domains.The oligonucleotides are also designed so
CC   that the expressed peptide contains 2 or 4 cysteine residues positioned
CC   in, or flanking, the unpredicted or variant residues. These residues
CC   confer some degree of conformational rigidity to the peptides. The TSARs
CC   or compsns. comprising a TSAR binding domain can be used in vivo to
CC   deliver a chemically or biologically active moiety, eg. metal ion,
CC   radioisotope, peptide, toxin or enzyme, to the specific target or on the
CC   cell. They can also replace the function of macromolecules, eg.
CC   monoclonal or polyclonal antibodies and therefore circumvent the need
CC   for complex methods of hybridoma formation or in vivo antibody
CC   production. The TSARs are easily characterised and have designed
CC   activity allowing direct and rapid detection in a screening process.
SQ   Sequence   114 BP;    0 A;    2 C;    2 G;    2 T;

  Query Match            1.8%;  Score 37;  DB 12;  Length 114;
  Best Local Similarity  5.5%;  Pred. No. 7.46e-08;
  Matches     6;  Conservative   32;  Mismatches  71;  Indels   0;  Gaps   0;

Db      6 bnnbnnbnnbnnbtgcnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnnnnnnnbnn 65
          :  :  :  :  :|||  :  :  :  :  :  :  :  :  :  :           :
Cp    515 CTCTGCCGTGGGCTGCAGGTGCTGCAACATGGTCTGGAACTCTGAGTCATAGCGTCGGTT 456

Db     66 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbtgcnnbnnbnnb 114
          :  :  :  :  :  :  :  :  :  :  :  :  :|||  :  :  :
Cp    455 GATGTCGTCCCCGATGATGGCGAGCTGCCGTCCCACCTGCCCCATGGTG 407


RESULT   14
ID   Q70468 standard; DNA; 114 BP.
AC   Q70468;
```

```
DT   05-APR-1995  (first entry)
DE   Generic DNA sequence to generate a random TSAR petide library.
KW   TSAR; totally synthetic affinity reagent; synthetic; binding domain;
KW   effector domain; concateneated heterofunctional protein; linker;
KW   direct; rapid; detection; screening; treatment; generic; ss.
OS   Synthetic.
FH   Key             Location/Qualifiers
FT   misc_feature    55..60
FT   /*tag= a
FT   /note= "this sequence represents 'Z'; Z can be a
FT   sequence of 6, 9 or 12 nucleotides (see
FT   comments)"
PN   WO9418318-A.
PD   18-AUG-1994.
PF   01-FEB-1994; U00977.
PR   01-FEB-1993; US-013416.
PR   30-DEC-1993; US-176500.
PR   31-JAN-1994; US-189331.
PA   (UYNC-) UNIV NORTH CAROLINA.
PI   Fowlkes DM,  Kay BK;
DR   WPI; 94-279739/34.
DR   P-PSDB; R65154.
PT   Identifying proteins or peptide(s) which bind a ligand - by
PT   screening a recombinant vector library expressing fusion proteins
PT   comprising a binding domain and an effector domain
PS   Disclosure; Page 35; 255pp; English.
CC   Q70468 is a generic DNA sequence used to generate random TSAR (Totally
CC   Synthetic Affinity Reagents) peptides.This generic formula can also be
CC   represented  as follows: X(NNB)11(TGC)(NNB)6Z(NNB)7(TGC)(NNB)10Y. X
CC   and Y are flanking  restriction sites (X is not the same as Y) that are
CC   not specified further. Other generic sequences are shown in Q70466-68.
CC   Other specific peptides generated by these generic sequences are shown in
CC   R65151-54. TSARs are concatenated heterofunctional proteins or peptides,
CC   comprising at least two functional regions - a binding domain with
CC   affinity for a ligand and a second effector peptide portion that is
CC   chemically or biologically active.They may further comprise a linker
CC   peptide between the 2 domains.The oligonucleotides are also designed so
CC   that the expressed peptide contains 2 or 4 cysteine residues positioned
CC   in, or flanking, the unpredicted or variant residues. These residues
CC   confer some degree of conformational rigidity to the peptides. The TSARs
CC   or compsns. comprising a TSAR binding domain can be used in vivo to
CC   deliver a chemically or biologically active moiety, eg. metal ion,
CC   radioisotope, peptide, toxin or enzyme, to the specific target or on the
CC   cell. They can also replace the function of macromolecules, eg.
CC   monoclonal or polyclonal antibodies and therefore circumvent the need
CC   for complex methods of hybridoma formation or in vivo antibody
CC   production. The TSARs are easily characterised and have designed activity
CC   allowing direct and rapid detection in a screening process.
SQ   Sequence   114 BP;    0 A;    2 C;    2 G;    2 T;

  Query Match            1.7%;  Score 36;  DB 12;  Length 114;
  Best Local Similarity  5.7%;  Pred. No. 2.86e-07;
  Matches     6;  Conservative   31;  Mismatches  69;  Indels   0;  Gaps   0;

Db      3 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbtgcnnbnnbnnbnnbnnbnnbnnnnnnnn 62
          :  :  :  :  :     :  :  :  :  :|||  :  :  :  :  :  :  :
Cp    129 GGAGATGGAGGTCCGAGGGGCTGGCGGCTGCTGCTCCCAGGGGCTGAGTGGGAGCCCAGT 70

Db     63 bnnbnnbnnbnnbnnbnnbtgcnnbnnbnnbnnbnnbnnbnnbnnb 108
           :  :  :  :  :  :  :|||  :  :  :  :  :  :  :  :
Cp     69 TTCCAGGAATGGGCGTCAGTGCATTCCCGGCATCTGGATGTAGCCT 24


RESULT   15
ID   Q70466 standard; DNA; 114 BP.
AC   Q70466;
DT   05-APR-1995  (first entry)
DE   Generic DNA sequence to generate a random TSAR-9 petide library.
KW   TSAR; totally synthetic affinity reagent; synthetic; binding domain;
KW   effector domain; concateneated heterofunctional protein; linker;
KW   direct; rapid; detection; screening; treatment; generic; ss.
OS   Synthetic.
FH   Key             Location/Qualifiers
FT   misc_feature    55..60
FT   /*tag= a
FT   /note= "this sequence represents 'Z'; Z can be a
FT   sequence of 6, 9 or 12 nucleotides (see
FT   comments)"
PN   WO9418318-A.
PD   18-AUG-1994.
PF   01-FEB-1994; U00977.
PR   01-FEB-1993; US-013416.
PR   30-DEC-1993; US-176500.
PR   31-JAN-1994; US-189331.
PA   (UYNC-) UNIV NORTH CAROLINA.
PI   Fowlkes DM,  Kay BK;
DR   WPI; 94-279739/34.
DR   P-PSDB; R65152.
PT   Identifying proteins or peptide(s) which bind a ligand - by
PT   screening a recombinant vector library expressing fusion proteins
PT   comprising a binding domain and an effector domain
PS   Disclosure; Page 35; 255pp; English.
CC   Q70466 is a generic DNA sequence used to generate random TSAR (Totally
CC   Synthetic Affinity Reagents) peptides.This generic formula can also be
CC   represented as follows:X(NNB)1(TGC)(NNB)10(TGC)2(NNB)4Z(NNB)8(TGC)(NNB)
CC   -9Y. X and Y are flanking  restriction sites (X is not the same as Y)
CC   that are not specified further. Other generic sequences are shown in
CC   Q70466-68. Other specific peptides generated by these generic sequences
CC   are shown in R65151-54. TSARs are concatenated heterofunctional proteins
CC   or peptides, comprising at least two functional regions - a binding
CC   domain with affinity for a ligand and a second effector peptide portion
CC   that is chemically or biologically active.They may further comprise a
CC   linker peptide between the 2 domains.The oligonucleotides are also
CC   designed so that the expressed peptide contains 2 or 4 cysteine residues
CC   positioned in, or flanking, the unpredicted or variant residues. These
CC   residues confer some degree of conformational rigidity to the peptides.
CC   The TSARs or compsns. comprising a TSAR binding domain can be used in
CC   vivo to deliver a chemically or biologically active moiety, eg. metal
CC   ion, radioisotope, peptide, toxin or enzyme, to the specific target or
CC   on the cell. They can also replace the function of macromolecules, eg.
CC   monoclonal or polyclonal antibodies and therefore circumvent the need for
CC   complex methods of hybridoma formation or in vivo antibody production.
CC   The TSARs are easily characterised and have designed activity allowing
CC   direct and rapid detection in a screening process.
SQ   Sequence   114 BP;    0 A;    4 C;    4 G;    4 T;

  Query Match            1.6%;  Score 34;  DB 12;  Length 114;
  Best Local Similarity  7.3%;  Pred. No. 4.05e-06;
  Matches     8;  Conservative   30;  Mismatches  72;  Indels   0;  Gaps   0;

Db      5 gcnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbtgctgcnnbnnbnnbnnbnnnnnnnnbn 64
          ||  :  :  :  :  :  :  :  :  :  :  :|||  |  :  :  :  :        :
Qy    565 GCATCAATTGGGGCCGTGTGGTGGCTCTTCTGGGCTTCGGCTACCGTCTGGCCCTACACG 624

Db     65 nbnnbnnbnnbnnbnnbnnbtgcnnbnnbnnbnnbnnbnnbnnbnnbnnb 114
           :  :  :  :  :  :  :  :||  :     :  :  :  :  :  :  :
Qy    625 TCTACCAGCATGGCCTGACTGGCTTCCTAGGCCAGGTGACCCGCTTCGTG 674


Search completed: Thu Aug 21 10:38:25 1997
Job time : 628 secs.
```

****

MPSRCH (TM)

****

Release 2.1D John F. Collins, Biocomputing Research Unit.

MPsrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Thu Aug 21 10:38:50 1997; MasPar time 1069.97 Seconds
1004.594 Million cell updates/sec

Tabular output not generated.

Title: >US-08-320-157-6
Description: (1-2094) from US08320157.seq
Perfect Score: 2094
N.A. Sequence: 1 GAGGATCTACAGGGGACAAG.....CAAAAAAAAAACGGAGATCC 2094
Comp: CTCCTAGATGTCCCCTGTTC.....GTTTTTTTTTTGCCTCTAGG

Scoring table: TABLE default
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 707517 seqs, 256659390 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: EST-STS
1:EST1 2:EST2 3:EST3 4:EST4 5:EST5 6:EST6 7:EST7 8:EST8
9:EST9 10:EST10 11:EST11 12:EST12 13:EST13 14:EST14
15:EST15 16:EST16 17:EST17 18:EST18 19:EST19 20:EST20
21:EST21 22:EST22 23:EST23 24:EST24 25:EST25 26:EST26
27:EST27 28:EST28 29:EST29 30:EST30 31:EST31 32:EST32
33:EST33 34:EST34 35:EST35 36:EST36 37:EST37 38:EST38
39:EST39 40:EST40 41:EST41 42:EST42 43:EST43 44:EST44
45:EST45 46:EST46 47:EST47 48:EST48 49:EST49 50:EST50
51:EST51 52:EST52 53:EST53 54:EST54 55:EST55 56:EST56
57:EST57 58:EST58 59:EST59 60:EST60 61:EST61 62:EST62
63:EST63 64:EST64 65:EST65 66:EST66 67:EST67 68:EST68
69:EST69 70:EST70 71:EST71 72:EST72 73:EST73 74:EST74
75:EST75 76:EST76 77:EST77 78:EST78 79:EST79 80:EST80
81:EST81 82:EST82 83:EST83 84:EST84 85:EST85 86:EST86
87:EST87 88:EST88 89:EST89 90:EST90 91:EST91 92:EST92
93:EST93 94:EST94 95:EST95 96:EST96 97:EST97 98:EST98
99:EST99
Database: EST-STS-TWO
100:EST100 101:EST101 102:EST102 103:EST103 104:EST104
105:EST105 106:EST106 107:EST107 108:EST108 109:EST109
110:EST110 111:EST111 112:EST112 113:EST113 114:EST114
115:EST115 116:EST116 117:EST117 118:EST118 119:EST119
120:EST120 121:EST121 122:EST122 123:EST123 124:EST124
125:EST125 126:EST126 127:EST127 128:EST128 129:EST129
130:EST130 131:EST131 132:EST132 133:EST133 134:EST134
135:EST135 136:EST136 137:EST137 138:EST138 139:EST139
140:EST140 141:EST141 142:EST142 143:EST143 144:EST144
145:EST145 146:EST146 147:EST147 148:EST148 149:EST149
150:EST150 151:EST151 152:EST152 153:EST153 154:EST154
155:EST155 156:EST156 157:EST157 158:EST158 159:EST159
160:EST160 161:EST161 162:EST162 163:EST163 164:EST164
165:EST165 166:EST166 167:EST167 168:EST168 169:EST169
170:EST170 171:EST171 172:EST172 173:EST173 174:EST174
175:EST175 176:EST176 177:EST177 178:EST178 179:EST179
180:EST180 181:EST181 182:EST182 183:EST183 184:EST184
185:EST185 186:EST186 187:EST187 188:EST188 189:EST189
190:EST190 191:EST191 192:EST192 193:EST193 194:EST194
195:EST195 196:EST196 197:EST197 198:EST198

Statistics: Mean 11.801; Variance 2.224; scale 5.307

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description	Pred. No.
	1	355	17.0	409	84	H52672	yu36g10.r1 Homo sapie	0.00e+00
c	2	339	16.2	449	84	H52673	yu36g10.s1 Homo sapie	0.00e+00
c	3	267	12.8	497	155	AA016399	mg88g02.r1 Soares mou	0.00e+00
	4	266	12.7	303	85	H56462	yt87a12.r1 Homo sapie	0.00e+00
c	5	248	11.8	545	153	W42014	mb16g04.r1 Soares mou	0.00e+00
c	6	229	10.9	446	138	AA049970	mj39b01.r1 Soares mou	0.00e+00
	7	154	7.4	161	151	W23886	zb46a10.r1 Soares fet	7.70e-255
	8	72	3.4	292	138	AA049828	mj09a09.r1 Soares mou	3.82e-86
	9	72	3.4	361	139	AA050569	mj16b12.r1 Soares mou	3.82e-86
	10	62	3.0	232	133	N83998	KK4271F Homo sapiens	4.39e-67
	11	46	2.2	270	175	W82028	mf02e01.r1 Soares mou	3.69e-38
	12	44	2.1	445	126	W77319	me64c04.r1 Soares mou	9.87e-35
	13	39	1.9	320	186	AA139013	mr04a06.r1 Soares mou	2.04e-26
c	14	33	1.6	408	82	H31839	EST106322 Rattus sp.	5.04e-17
	15	31	1.5	279	82	H31840	EST106323 Rattus sp.	4.50e-14
c	16	25	1.2	182	99	N62399	yz73c08.s1 Homo sapie	6.17e-06
	17	25	1.2	386	2	T58949	yb81f10.r1 Homo sapie	6.17e-06
c	18	26	1.2	480	5	T70889	yc49b08.s1 Homo sapie	3.32e-07
c	19	25	1.2	500	61	H15311	ym28b07.r1 Homo sapie	6.17e-06
c	20	23	1.1	219	22	R39042	yd07g10.s1 Homo sapie	1.59e-03
c	21	23	1.1	309	140	C03863	Human Heart cDNA, clo	1.59e-03
c	22	24	1.1	337	109	HSC3LC012	H. sapiens partial cD	1.04e-04
c	23	23	1.1	337	145	R85722	yt67d12.s1 Soares ret	1.59e-03
c	24	23	1.1	338	105	HSB73E022	H. sapiens partial cD	1.59e-03
c	25	24	1.1	358	133	N92010	za21a11.r1 Homo sapie	1.04e-04
	26	23	1.1	368	194	AA169681	zo92f06.r1 Stratagene	1.59e-03
	27	23	1.1	368	40	R22823	yh20h06.s1 Homo sapie	1.59e-03
c	28	23	1.1	377	19	T54824	yb43f08.r1 Homo sapie	1.59e-03
c	29	23	1.1	391	115	W17230	zb14a04.r1 Soares fet	1.59e-03
	30	23	1.1	405	113	W02475	za47d08.r1 Soares fet	1.59e-03
	31	23	1.1	406	136	AA001141	zh83c07.r1 Soares fet	1.59e-03
c	32	24	1.1	409	175	W84628	zd91f06.s1 Soares fet	1.04e-04
c	33	23	1.1	412	159	AA036957	zk32a09.s1 Soares pre	1.59e-03
	34	23	1.1	423	188	AA146773	zo35f12.s1 Stratagene	1.59e-03
c	35	24	1.1	425	147	AA004647	zh92c08.s1 Soares fet	1.04e-04
	36	24	1.1	430	41	R25797	yg54b04.r1 Homo sapie	1.04e-04
	37	24	1.1	432	53	R92375	yq06g09.r1 Homo sapie	1.04e-04
c	38	23	1.1	445	174	W80365	zd82a03.s1 Soares fet	1.59e-03
c	39	24	1.1	446	170	W57984	zd18h07.s1 Soares fet	1.04e-04
c	40	23	1.1	456	49	HUM519C08B	Human placenta cDNA 5	1.59e-03
c	41	23	1.1	465	61	H16794	ym39b06.s1 Homo sapie	1.59e-03
	42	24	1.1	468	147	AA004387	zh92d01.r1 Soares fet	1.04e-04
c	43	24	1.1	480	136	AA001756	zh84b09.s1 Soares fet	1.04e-04
c	44	24	1.1	487	125	W74617	zd77e01.s1 Soares fet	1.04e-04
c	45	23	1.1	641	183	AA097231	mk10f11.r1 Soares mou	1.59e-03

ALIGNMENTS

RESULT 1
LOCUS H52672 409 bp mRNA EST 18-SEP-1995
DEFINITION yu36g10.r1 Homo sapiens cDNA clone 235938 5'.
ACCESSION H52672
NID g992513
KEYWORDS EST.
SOURCE clone=235938 primer=M13RP1 library=Soares ovary tumor NbHOT

```
            vector=pT7T3D (Pharmacia) with a modified polylinker host=DH10B
            (ampicillin resistant) Rsite1=Not I Rsite2=Eco RI Female. 1st
            strand cDNA was primed with a Not I - oligo(dT) primer [5'
            TGTTACCAATCTGAAGTGGGAGCGGCCGCGGTTTTTTTTTTTTTTTTT 3'],
            double-stranded cDNA was size selected, ligated to Eco RI adapters
            (Pharmacia), digested with Not I and cloned into the Not I and Eco
            RI sites of a modified pT7T3 vector (Pharmacia). Library
            constructed by Bento Soares and M.Fatima Bonaldo.
  ORGANISM  Homo sapiens
            Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata;
            Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
            Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria;
            Eutheria; Archonta; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 409)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT
            Contact: Wilson RK
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            High quality sequence stops: 307
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
FEATURES              Location/Qualifiers
     source           1..409
                      /organism="Homo sapiens"
                      /clone="235938"
     mRNA             <1..>409
BASE COUNT       73 a    127 c    101 g    101 t      7 others
ORIGIN

  Query Match           17.0%;  Score 355;  DB 84;  Length 409;
  Best Local Similarity  96.1%;  Pred. No. 0.00e+00;
  Matches   390;  Conservative    0;  Mismatches  11;  Indels   5;  Gaps   5;

Db      1 ctatgttccccaggattcagctattctggaagatcagcaccctaagagatgggactagga 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1263 CTATGTTCCCCAGGATTCAGCTATTCTGGAAGATCAGCACCCTAAGAGATGGGACTAGGA 1322

Db     61 cctgagcctggtcctggccgtccctaagcatgtgtcccaggagcaggacctactaggaga 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1323 CCTGAGCCTGGTCCTGGCCGTCCCTAAGCATGTGTCCCAGGAGCAGGACCTACTAGGAGA 1382

Db    121 ggggggccaaggtcctgctcaactctacccctgctcccattcctccctccggccatactg 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1383 GGGGGGCCAAGGTCCTGCTCAACTCTACCCCTGCTCCCATTCCTCCCTCCGGCCATACTG 1442

Db    181 cctttgcagttggactctcagggattctgggcttggggtgtggggtngggtngagtcgca 240
          ||||||||||||||||||||||||||||||||||||||||||||||||| |||| ||||||||
Qy   1443 CCTTTGCAGTTGGACTCTCAGGGATTCTGGGCTTGGGGTGTGGGGTGGGGTGGAGTCGCA 1502

Db    241 gac-agagctgtctgaactcacgtgtcagaagcctccaagcctnccctcccaaggtcctct 299
          ||| |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Qy   1503 GACCAGAGCTGTCTGAACTCACGTGTCAGAAGCCTCCAAGCCTGCCTCCCAAGGTCCTCT 1562

Db    300 cagttctctccccttcctctctccttatagacacttgctcccaacccattncactacaggt 359
          |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Qy   1563 CAGTTCTCTCCCTTCCTCTCTCCTTATAGACACTTGCTCCCAACCCATT-CACTACAGGT 1621

Db    360 naaggcttttcacccccatcctggggnccttnggttgagttgcctg 405
           |||||| | |||||    ||||||| |||| ||| |||| |||||
Qy   1622 GAAGGCTCT-CACCCATC-CCTGGGGGCCTTGGGT-GAGTGGCCTG 1664
```

```
RESULT    2
LOCUS       H52673        449 bp    mRNA            EST       18-SEP-1995
DEFINITION  yu36g10.s1 Homo sapiens cDNA clone 235938 3'.
ACCESSION   H52673
NID         g992514
KEYWORDS    EST.
SOURCE       clone=235938 primer=Promega -21m13 library=Soares ovary tumor
            NbHOT vector=pT7T3D (Pharmacia) with a modified polylinker
            host=DH10B (ampicillin resistant) Rsite1=Not I Rsite2=Eco RI
            Female. 1st strand cDNA was primed with a Not I - oligo(dT) primer
            [5' TGTTACCAATCTGAAGTGGGAGCGGCCGCGGTTTTTTTTTTTTTTTTT 3'],
            double-stranded cDNA was size selected, ligated to Eco RI adapters
            (Pharmacia), digested with Not I and cloned into the Not I and Eco
            RI sites of a modified pT7T3 vector (Pharmacia). Library
            constructed by Bento Soares and M.Fatima Bonaldo.
  ORGANISM  Homo sapiens
            Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata;
            Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
            Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria;
            Eutheria; Archonta; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 449)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT
            Contact: Wilson RK
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            High quality sequence stops: 379
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
FEATURES              Location/Qualifiers
     source           1..449
                      /organism="Homo sapiens"
                      /clone="235938"
     mRNA             <1..>449
BASE COUNT      106 a    127 c    113 g     95 t      8 others
ORIGIN

  Query Match           16.2%;  Score 339;  DB 84;  Length 449;
  Best Local Similarity  94.5%;  Pred. No. 0.00e+00;
  Matches   413;  Conservative    0;  Mismatches  13;  Indels  11;  Gaps  11;

Db     11 attncacagtttatttccaaacactcagaggganagggggtggcctatggg-tccatctgc 69
          ||| |||||||||||||||||||||||| |||||||||||||||| |||||||||
Cp   2069 ATTGCACAGTTTATTTCCAAACACTCAGAGGATAGGGGGTGGCCTATGGGCTCCATCTGC 2010

Db     70 ctcccctccctgcatttggctgaatcaagaacttctccccccctgacccccaagaccctag 129
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Cp   2009 CTCCCCTCCCTGCATTTGGCTGAATCAAGAACTTCTCCCCCCTGACCCCCAAGACCCTAG 1950

Db    130 gctgtgccaatagagaaggcactgtcacccccccaccctgaggagacctagacaggtgagg 189
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Cp   1949 GCTGTGCCAATAGAGAAGGCACTGTCACCCCCCACCCTGAGGAGACCTAGACAGGTGAGG 1890

Db    190 agcatggagggtgggggaacagagttcaagtattcacagttccccccatctacaccccggg 249
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Cp   1889 AGCATGGAGGGTGGGGGAACAGAGTTCAAGTATTCACAGTTCCCCCATCTACACCCCTGG 1830

Db    250 attacactgtgccagagccatgagggaggatcccacctctgggattcctagtgggtgttg 309
          |||||||||||||||||||||||||||||||| |||||||||||||||||| |||||
Cp   1829 ATTACACTGTGCCAGAGCCATGAGGGAGGATCC-ACCTCTGGGATTCCTAGTGG-TGTTG 1772
```

```
Db    310 atagtccttctcccacttaggaaccctccagatgaactccctactccttttccctggaaa 369
          ||||||||||||||||||| |||||||||||||||||||||||||||||||||| | |
Cp   1771 ATAGTCCTTCTCCCACTTAG-AACCCTCCAGATGAACTCCCTACTCCTTTTCCCTG-ATA 1714

Db    370 taacaaaccaaggtcctaaggccctgtagtttggggcaagggagggagtntaggnagggc 429
          |||||||||||| ||||||| ||||||||| |||| ||||| ||| ||  |  | ||| |
Cp   1713 TAACAAACCAAG-TCCTAAG-CCCTGTAGTCTGGG-CAAGG-AGG-AGCCTTAGCAGG-C 1660

Db    430 cattnacccaagggccc 446
          || | ||||||||| |||
Cp   1659 CACTCACCCAAGGCCCC 1643


RESULT    3
LOCUS       AA016399      497 bp    mRNA            EST       02-AUG-1996
DEFINITION  mg88g02.r1 Soares mouse embryo NbME13.5 14.5 Mus musculus cDNA
            clone 440114 5' similar to SW:BCLX_HUMAN Q07817 APOPTOSIS REGULATOR
            BCL-X. ;.
ACCESSION   AA016399
NID         g1478767
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae;
            Murinae; Mus.
REFERENCE   1  (bases 1 to 497)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT
            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:265450
            Possible reversed clone: similarity on wrong strand
            Seq primer: -28M13 rev2 from Amersham
            High quality sequence stop: 484.
FEATURES              Location/Qualifiers
     source           1..497
                      /organism="Mus musculus"
                      /strain="C57BL/6J"
                      /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                      polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                      was primed with a Not I - oligo(dT) primer [5'
                      TGTTACCAATCTGAAGTGGGAGCGGCCGCGGAAATTTTTTTTTTTTTTTTTTTTTT
                      T 3'], on equal amounts of mRNA from 2 13.5dpc and 2
                      14.5dpc embryos [total RNA provided by  Minoru Ko, Wayne
                      State Univ., from 2 ]; double-stranded cDNA was ligated to
                      Eco RI adaptors (Pharmacia), digested with Not I and
                      cloned into the Not I and Eco RI sites of the modified
                      pT7T3 vector. Library went through one round of
                      normalization, and was constructed by Bento Soares and
                      M.Fatima Bonaldo."
                      /clone="440114"
                      /clone_lib="Soares mouse embryo NbME13.5 14.5"
                      /sex="unknown"
                      /tissue_type="embryo"
                      /dev_stage="13.5-14.5dpc total fetus"
                      /lab_host="DH10B"
     mRNA             <1..>497
BASE COUNT      126 a    137 c    137 g     97 t
ORIGIN
```

```
  Query Match           12.8%;  Score 267;  DB 155;  Length 497;
  Best Local Similarity  78.1%;  Pred. No. 0.00e+00;
  Matches   371;  Conservative    0;  Mismatches 104;  Indels   0;  Gaps   0;

Db     22 tttagtccagccaggcacgtggctgggtgccccagatcttcacccgggagtcatgatctg 81
          || ||||||| ||||    ||    ||  ||| |      | |  |||||||||||| ||
Cp    886 TTAAGTCCAGGCAGGGGTCTGAACCGGGACCCAAAGGGCACCCTTGGGAGTCATGATTTG 827

Db     82 aagaatctgtgtaccacgaattggcccaacagaaccacaccaaaaatcaccattacggtc 141
          ||||||||  ||||||| || ||||||||||||||||||| | || |||||  ||| ||
Cp    826 AAGAATCTTCGTACCACAAACTGGCCCAACAGAACCACACCCAGAACCACCAGCACGTTC 767

Db    142 aggatggggtctctacgaaaattcagggctgccacccaaccgcctctctgtgcgatccat 201
          ||||||||  |  | |  || ||||||||||||||||| || || |||||||| |||||
Cp    766 AGGATGGGACCATTGCCCAAGTTCAGGGCTGCCACCCAGCCACCCCTCTGTGCAATCCAC 707

Db    202 ctggcgatgtaatgatgcagtatgatatcagccaaaaagcaggtcacctggcccaggaag 261
          | ||| ||| | |||||||| |||| || ||| |||| |||||||||||| ||||||
Cp    706 CGGGCAATGCAGTGATGCAGCATGAAGTCGACCACGAAGCGGGTCACCTGGCCTAGGAAG 647

Db    262 ccggtcaaaccacgctggtagacgtacagggccagacggtagccaaagcccaggagagcc 321
          || ||||  ||| ||||||||||||  ||||||||||||||||| ||||||||  ||||||
Cp    646 CCAGTCAGGCCATGCTGGTAGACGTGTAGGGCCAGACGGTAGCCGAAGCCCAGAAGAGCC 587

Db    322 accacgcggccccagctgatgccactcttaaataggctggaggcgatcttggtgaagagt 381
          ||||| ||||||||  |||||||||||| ||| ||||||| ||| ||||||||||||
Cp    586 ACCACACGGCCCCAATTGATGCCACTCTCAAACAGGCTGGTGGCAATCTTGGTGAAGTAC 527

Db    382 tcgtaggcattcccggctgtgggctgaagctgttctagtaaattctggaactctgtgtcg 441
          || ||||||||| | || |||||||| || || |  |  |   |||||||||||| |||
Cp    526 TCATAGGCATTCTCTGCCGTGGGCTGCAGGTGCTGCAACATGGTCTGGAACTCTGAGTCA 467

Db    442 tagcgccggttaatatcatctccgatgagagcaagctgccgacccacctgactca 496
          ||||| ||||| || || || |||||||  || |||||||| |||||||| | ||
Cp    466 TAGCGTCGGTTGATGTCGTCCCCGATGATGGCGAGCTGCCGTCCCACCTGCCCCA 412


RESULT    4
LOCUS       H56462        303 bp    mRNA            EST       02-OCT-1995
DEFINITION  yt87a12.r1 Homo sapiens cDNA clone 231262 5'.
ACCESSION   H56462
NID         g1005106
KEYWORDS    EST.
SOURCE       clone=231262 primer=M13RP1 library=Soares pineal gland N3HPG
            vector=pT7T3D (Pharmacia) with a modified polylinker host=DH10B
            (ampicillin resistant) Rsite1=Not I Rsite2=Eco RI 1st strand cDNA
            was primed with a Not I - oligo(dT) primer [5'
            TGTTACCAATCTGAAGTGGGAGCGGCCGCGTTTTTTTTTTTTTTTTTT 3'],
            double-stranded cDNA was size selected, ligated to Eco RI adapters
            (Pharmacia), digested with Not I and cloned into the Not I and Eco
            RI sites of a modified pT7T3 vector (Pharmacia). Library
            constructed by Bento Soares and M.Fatima Bonaldo.
  ORGANISM  Homo sapiens
            Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata;
            Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
            Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria;
            Eutheria; Archonta; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 303)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT
            Contact: Wilson RK
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
```

```
            Email: est@watson.wustl.edu
            High quality sequence stops: 251
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
FEATURES              Location/Qualifiers
     source           1..303
                      /organism="Homo sapiens"
                      /clone="231262"
     mRNA             <1..>303
BASE COUNT       68 a     91 c     90 g     50 t      4 others
ORIGIN

  Query Match           12.7%;  Score 266;  DB 85;  Length 303;
  Best Local Similarity  94.7%;  Pred. No. 0.00e+00;
  Matches   288;  Conservative    0;  Mismatches  14;  Indels   2;  Gaps   2;

Db      1 gggcaaggcaatgnnccctcccaggcaggagtgcggagagcctgccctgccctctgcttct 60
          |||||||||   |  ||||||||||||||||||||||||||||||||||||||||||||||
Qy    210 GGGCAAGGCCCAGGTCCTCCCAGGCAGGAGTGCGGAGAGCCTGCCCTGCCCTCTGCTTCT 269

Db     61 gaggagcaggtagcccaggacacagaggaggttttccgcagctacgttttttaccaccat 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Qy    270 GAGGAGCAGGTAGCCCAGGACACAGAGGAGGTTTTCCGCAGCTACGTTTTTTACCGCCAT 329

Db    121 cagcaggaacaggaggctgaaggggcggctgctgacnc-gacccagagatggtcaccttа 179
          |||||||||||||||||||||||| ||||||     | ||||||||||||||||||||
Qy    330 CAGCAGGAACAGGAGGCTGAAGGGGTGGCTGCCCCTGCCGACCCAGAGATGGTCACCTTA 389

Db    180 cctctgcaacctagcagcaccatggggcaggtgggacggcagctcgccatcattggggac 239
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Qy    390 CCTCTGCAACCTAGCAGCACCATGGGGCAGGTGGGACGGCAGCTCGCCATCATCGGGGAC 449

Db    240 gacatcaaccgacgctatgactcagagttccagaccatgttncagcacctgcagtcccac 299
          |||||||||||||||||||||||||||||||||||||||| ||||||||||| |||||
Qy    450 GACATCAACCGACGCTATGACTCAGAGTTCCAGACCATGTTGCAGCACCTGCAG-CCCAC 508

Db    300 ggca 303
          ||||
Qy    509 GGCA 512


RESULT    5
LOCUS       W42014        545 bp    mRNA            EST       11-SEP-1996
DEFINITION  mb16g04.r1 Soares mouse p3NMF19.5 Mus musculus cDNA clone 329622 5'
            similar to SW:BCLX_HUMAN Q07817 APOPTOSIS REGULATOR BCL-X. ;.
ACCESSION   W42014
NID         g1326554
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Mus.
REFERENCE   1  (bases 1 to 545)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT
            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:211022

            Possible reversed clone: similarity on wrong strand
            Seq primer: ETPrimer
            High quality sequence stop: 411.
FEATURES              Location/Qualifiers
     source           1..545
                      /organism="Mus musculus"
                      /note="Vector: pT7T3D (Pharmacia) with a modified
                      polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                      was primed with a Not I - oligo(dT) primer [5'
                      TGTTACCAATCTGAAGTGGGAGCGGCCGCATTTTTTTTTTTTTTTTTTT 3'],
                      double-stranded cDNA was size selected, ligated to Eco RI
                      adapters (Pharmacia), digested with Not I and cloned into
                      the Not I and Eco RI sites of a modified pT7T3 vector
                      (Pharmacia). Library went through one round of
                      normalization to a Cot = 5. Library constructed by Bento
                      Soares and M.Fatima Bonaldo. RNA was kindly provided by
                      Dr. Minoru Ko (Wayne State University)."
                      /clone="329622"
                      /clone_lib="Soares mouse p3NMF19.5"
                      /dev_stage="19.5 dpc total fetus"
                      /lab_host="DH10B (ampicillin resistant)"
     mRNA             <1..>545
BASE COUNT      134 a    139 c    165 g    105 t      2 others
ORIGIN

  Query Match           11.8%;  Score 248;  DB 153;  Length 545;
  Best Local Similarity  77.9%;  Pred. No. 0.00e+00;
  Matches   381;  Conservative    0;  Mismatches 103;  Indels   5;  Gaps   4;

Db     31 tttagtccagccaggcacgtggctggggccccagtgggtcccc--gggagtcatgatctg 88
          || ||||||| ||||    ||    ||| ||||  |||  |||  ||||||||||||| ||
Cp    886 TTAAGTCCAGGCAGGGGTCTGAACCGGGACCCAAAGGGCACCCTTGGGAGTCATGATTTG 827

Db     89 aagaatctgtgtaccacgaattggcccaacagaaccacaccaaaaatcaccattacggtc 148
          ||||||||  ||||||| || ||||||||||||||||||| | || |||||  ||| ||
Cp    826 AAGAATCTTCGTACCACAAACTGGCCCAACAGAACCACACCCAGAACCACCAGCACGTTC 767

Db    149 aggatggggactctacgaaaattcagggctgccacccaaccgcctctctgtgcgatccat 208
          ||||||||  |  | |  || |||||||||||||||| || || |||||||| |||||
Cp    766 AGGATGGGACCATTGCCCAAGTTCAGGGCTGCCACCCAGCCACCCCTCTGTGCAATCCAC 707

Db    209 ctggcgatgtaatgatgcagtatgatatcagccaaaaagcaggtcacctggcccaggaag 268
          | ||| ||| | |||||||| ||||  ||  |||  |||| ||||||||||||| ||||||
Cp    706 CGGGCAATGCAGTGATGCAGCATGAAGTCGACCACGAAGCGGGTCACCTGGCCTAGGAAG 647

Db    269 ccggtcaaaccacgctggtagacgtacagggccagacggtagccaaagcccaggagagcc 328
          || ||||  ||| |||||||||||  ||||||||||||||||| ||||||| ||||||
Cp    646 CCAGTCAGGCCATGCTGGTAGACGTGTAGGGCCAGACGGTAGCCGAAGCCCAGAAGAGCC 587

Db    329 accacgcggccccagctgatgccactcttaaataggctggaggcgatcttggtgaagagt 388
          ||||| ||||||||  |||||||||||| ||| ||||||| ||| ||||||||||||
Cp    586 ACCACACGGCCCCAATTGATGCCACTCTCAAACAGGCTGGTGGCAATCTTGGTGAAGTAC 527

Db    389 tcgtaggcattcccggctgtgggctgaagctgttctagtagattctggaactctgtgtcg 448
          || ||||||||| | || |||||||| || || |  |  |   |||||||||||| |||
Cp    526 TCATAGGCATTCTCTGCCGTGGGCTGCAGGTGCTGCAACATGGTCTGGAACTCTGAGTCA 467

Db    449 ta-cgccggttaatatca-ctccgatgagagcaagctgccgacc-acgttgaccaagatg 505
          || || ||||| || ||  | |||||||  || |||||||| || || |   ||| | ||
Cp    466 TAGCGTCGGTTGATGTCGTCCCCGATGATGGCGAGCTGCCGTCCCACCTGCCCCATGGTG 407

Db    506 ctgttgggt 514
          ||| | |||
Cp    406 CTGCTAGGT 398


RESULT    6
LOCUS       AA049970      446 bp    mRNA            EST       09-SEP-1996
DEFINITION  mj39b01.r1 Soares mouse embryo NbME13.5 14.5 Mus musculus cDNA
            clone 478441 5' similar to SW:BCLX_HUMAN Q07817 APOPTOSIS REGULATOR
            BCL-X. ;.
ACCESSION   AA049970
NID         g1529641
```

```
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Mus.
REFERENCE   1  (bases 1 to 446)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT
            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:289185
            Possible reversed clone: similarity on wrong strand
            Seq primer: -28M13 rev2 from Amersham
            High quality sequence stop: 319.
FEATURES              Location/Qualifiers
     source           1..446
                      /organism="Mus musculus"
                      /strain="C57BL/6J"
                      /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                      polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                      was primed with a Not I - oligo(dT) primer [5'
                      TGTTACCAATCTGAAGTGGGAGCGGCCGCGGAAATTTTTTTTTTTTTTTTTTTTTT
                      T 3'], on equal amounts of mRNA from 2 13.5dpc and 2
                      14.5dpc embryos [total RNA provided by  Minoru Ko, Wayne
                      State Univ., from 2 ]; double-stranded cDNA was ligated to
                      Eco RI adaptors (Pharmacia), digested with Not I and
                      cloned into the Not I and Eco RI sites of the modified
                      pT7T3 vector. Library went through one round of
                      normalization, and was constructed by Bento Soares and
                      M.Fatima Bonaldo."
                      /clone="478441"
                      /clone_lib="Soares mouse embryo NbME13.5 14.5"
                      /sex="unknown"
                      /tissue_type="embryo"
                      /dev_stage="13.5-14.5dpc total fetus"
                      /lab_host="DH10B"
     mRNA             <1..>446
BASE COUNT      110 a    118 c    127 g     91 t
ORIGIN
```

```
  Query Match           10.9%;  Score 229;  DB 138;  Length 446;
  Best Local Similarity  79.3%;  Pred. No. 0.00e+00;
  Matches   310;  Conservative    0;  Mismatches  81;  Indels   0;  Gaps   0;

Db     56 agtgcgtccccggttgtcatgatctgaagaatctgtgtaccacgaattggcccaacagaa 115
          ||  ||  ||  ||  |||||||| ||||||||||  ||||||| || ||||||||||||
Cp    852 AGGGCACCCTTGGGAGTCATGATTTGAAGAATCTTCGTACCACAAACTGGCCCAACAGAA 793

Db    116 ccacaccaaaaatcaccattacggtcatgatggggtctctacgaaaattcagggctgcca 175
          |||||||  | ||  |||||  |||  ||| ||||||  |  | |  ||  ||||||||||||
Cp    792 CCACACCCAGAACCACCAGCACGTTCAGGATGGGACCATTGCCCAAGTTCAGGGCTGCCA 733

Db    176 cccaaccgcctctctgtgcgatccatctggcgatgtaatgatgcagtatgatatcagcca 235
          ||||  ||  ||  ||||||||  |||||  |  |||  |||  |  ||||||||  ||||   ||  |||
Cp    732 CCCAGCCACCCCTCTGTGCAATCCACCGGGCAATGCAGTGATGCAGCATGAAGTCGACCA 673

Db    236 aaaagcaggtcacctggcccaggaagccggtcaaaccacgctggtagacgtacagggcca 295
            ||||  ||||||||||||  ||||||||  ||||   |||  ||||||||||||   |||||||
Cp    672 CGAAGCGGGTCACCTGGCCTAGGAAGCCAGTCAGGCCATGCTGGTAGACGTGTAGGGCCA 613

Db    296 gacggtagccaaagcccaggagagccaccacgcggccccagctgatgccactcttaaaata 355
          ||||||||||  ||||||||  ||||||||||  ||||||||   ||||||||||||  |||  |
Cp    612 GACGGTAGCCGAAGCCCAGAAGAGCCACCACACGGCCCCAATTGATGCCACTCTCAAACA 553

Db    356 ggctggaggcgatcttggtgaagagttcgtaggcatttccggctgtgggctgaagctgtt 415
          ||||||  |||  ||||||||||||    ||  ||||||||   |  ||  ||||||||  ||  ||  |
Cp    552 GGCTGGTGGCAATCTTGGTGAAGTACTCATAGGCATTCTCTGCCGTGGGCTGCAGGTGCT 493

Db    416 ctagtaaattcttgaactctgtgtcgtagcg 446
            |   |    |||  ||||||||  |||  |||||
Cp    492 GCAACATGGTCTGGAACTCTGAGTCATAGCG 462
```

```
RESULT    7
LOCUS       W23886        161 bp    mRNA            EST       20-AUG-1996
DEFINITION  zb46a10.r1 Soares fetal lung NbHL19W Homo sapiens cDNA clone 306618
            5'.
ACCESSION   W23886
NID         g1300701
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 161)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Tan,F., Trevaskis,E.,
            Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R.
  TITLE     WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT
            Contact: Wilson RK
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 988   Std Error: 0.00
            Seq primer: mob.REGA+ET.
FEATURES              Location/Qualifiers
     source           1..161
                      /organism="Homo sapiens"
                      /note="Organ: lung; Vector: pT7T3D (Pharmacia) with a
                      modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                      strand cDNA was primed with a Not I - oligo(dT) primer
                      [5'-TGTTACCAATCTGAAGTGGGAGCGGCCGCAATTTTTTTTTTTTTTTTT-3'],
                      double-stranded cDNA was size selected, ligated to Eco RI
                      adapters (Pharmacia), digested with Not I and cloned into
                      the Not I and Eco RI sites of a modified pT7T3 vector
                      (Pharmacia). Library went through one round of
                      normalization to a Cot = 5. Library constructed by Bento
                      Soares and M.Fatima Bonaldo. This library was constructed
                      from the same fetus as the fetal heart library, Soares
                      fetal heart NbHH19W."
                      /clone="306618"
                      /clone_lib="Soares fetal lung NbHL19W"
                      /dev_stage="19 weeks"
                      /lab_host="DH10B (ampicillin resistant)"
     mRNA             <1..>161
BASE COUNT       33 a     46 c     38 g     44 t
ORIGIN
```

```
  Query Match           7.4%;  Score 154;  DB 151;  Length 161;
  Best Local Similarity  99.4%;  Pred. No. 7.70e-255;
  Matches   160;  Conservative    0;  Mismatches   0;  Indels   1;  Gaps   1;

Db      1 cctccctcagctctctgggacctccttagccctgtctgctaggcgctggggagactgata 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

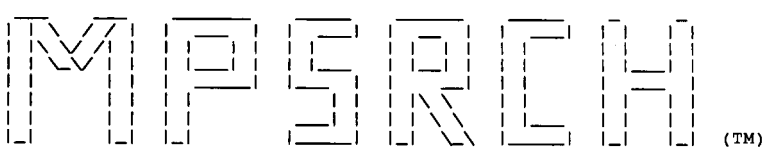
```
Qy   1063 CCTCCCTCAGCTCTCTGGGACCTCCTTAGCCCTGTCTGCTAGGCGCTGGGGAGACTGATA 1122

Db     61 acttggggaggcaagagactgggagccacttctccccagaaagtgtttaacggttttagc 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1123 ACTTGGGGAGGCAAGAGACTGGGAGCCACTTCTCCCCAGAAAGTGTTTAACGGTTTTAGC 1182

Db    121 tttttataataccttgtgagaagcccattcccaccattct 161
          ||||||||||||||||||||| ||||||||||||||||||
Qy   1183 TTTTTATAATACCCTTGTGAGA-GCCCATTCCCACCATTCT 1222


RESULT    8
LOCUS       AA049828      292 bp    mRNA            EST       09-SEP-1996
DEFINITION  mj09a09.r1 Soares mouse embryo NbME13.5 14.5 Mus musculus cDNA
            clone 475576 5'.
ACCESSION   AA049828
NID         g1529499
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Mus.
REFERENCE   1  (bases 1 to 292)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT
            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:286320
            Seq primer: -28M13 rev2 from Amersham.
FEATURES              Location/Qualifiers
     source           1..292
                      /organism="Mus musculus"
                      /strain="C57BL/6J"
                      /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                      polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                      was primed with a Not I - oligo(dT) primer [5'
                      TGTTACCAATCTGAAGTGGGAGCGGCCGCGGAAATTTTTTTTTTTTTTTTTTTTTT
                      T 3'], on equal amounts of mRNA from 2 13.5dpc and 2
                      14.5dpc embryos [total RNA provided by  Minoru Ko, Wayne
                      State Univ., from 2 ]; double-stranded cDNA was ligated to
                      Eco RI adaptors (Pharmacia), digested with Not I and
                      cloned into the Not I and Eco RI sites of the modified
                      pT7T3 vector. Library went through one round of
                      normalization, and was constructed by Bento Soares and
                      M.Fatima Bonaldo."
                      /clone="475576"
                      /clone_lib="Soares mouse embryo NbME13.5 14.5"
                      /sex="unknown"
                      /tissue_type="embryo"
                      /dev_stage="13.5-14.5dpc total fetus"
                      /lab_host="DH10B"
     mRNA             <1..>292
BASE COUNT       67 a     88 c     83 g     54 t
ORIGIN

  Query Match            3.4%;  Score 72;  DB 138;  Length 292;
  Best Local Similarity  79.7%;  Pred. No. 3.82e-86;
  Matches   137;  Conservative    0;  Mismatches  29;  Indels   6;  Gaps   4;

Db    127 ctccaccaagacctgaaaaatggcatctggacaaggaccaggtcccccgaaggtgggctg 186
          ||||||  ||||||||||||||| || || ||||| |||||||| || | |  ||  ||
Qy    182 CTCCACTGAGACCTGAAAAATGGCTTCGGGGCAAGGCCCAGGTCCTCCCAGGCAGGAGTG 241

Db    187 cgatgagtc--ccc-gtcc-c--cttctgaacagcaggttgcccaggacacagaggaggt 240
          ||  ||| |  ||| | || |  |||||||  ||||||| ||||||||||||||||||||
Qy    242 CGGAGAGCCTGCCCTGCCCTCTGCTTCTGAGGAGCAGGTAGCCCAGGACACAGAGGAGGT 301

Db    241 ctttcgaagctacgtttttttacctccaccagcaggaacaggagacccagggg 292
          || || ||||||||||||||| ||| |||||||||||||| |  | |||
Qy    302 TTTCCGCAGCTACGTTTTTTACCGCCATCAGCAGGAACAGGAGGCTGAAGGG 353


RESULT    9
LOCUS       AA050569      361 bp    mRNA            EST       09-SEP-1996
DEFINITION  mj16b12.r1 Soares mouse embryo NbME13.5 14.5 Mus musculus cDNA
            clone 476255 5'.
ACCESSION   AA050569
NID         g1530241
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Mus.
REFERENCE   1  (bases 1 to 361)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT
            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:286999
            Seq primer: -28M13 rev2 from Amersham.
FEATURES              Location/Qualifiers
     source           1..361
                      /organism="Mus musculus"
                      /strain="C57BL/6J"
                      /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                      polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                      was primed with a Not I - oligo(dT) primer [5'
                      TGTTACCAATCTGAAGTGGGAGCGGCCGCGGAAATTTTTTTTTTTTTTTTTTTTTT
                      T 3'], on equal amounts of mRNA from 2 13.5dpc and 2
                      14.5dpc embryos [total RNA provided by  Minoru Ko, Wayne
                      State Univ., from 2 ]; double-stranded cDNA was ligated to
                      Eco RI adaptors (Pharmacia), digested with Not I and
                      cloned into the Not I and Eco RI sites of the modified
                      pT7T3 vector. Library went through one round of
                      normalization, and was constructed by Bento Soares and
                      M.Fatima Bonaldo."
                      /clone="476255"
                      /clone_lib="Soares mouse embryo NbME13.5 14.5"
                      /sex="unknown"
                      /tissue_type="embryo"
                      /dev_stage="13.5-14.5dpc total fetus"
                      /lab_host="DH10B"
     mRNA             <1..>361
BASE COUNT       87 a    108 c    103 g     63 t
ORIGIN

  Query Match            3.4%;  Score 72;  DB 139;  Length 361;
  Best Local Similarity  79.7%;  Pred. No. 3.82e-86;
```

```
  Matches   137;  Conservative    0;  Mismatches  29;  Indels   6;  Gaps   4;

Db    196 ctccaccaagacctgaaaaatggcatctggacaaggaccaggtcccccgaaggtgggctg 255
          ||||||  ||||||||||||||| || || ||||| ||||||| || | |   ||  ||
Qy    182 CTCCACTGAGACCTGAAAAATGGCTTCGGGGCAAGGCCCAGGTCCTCCCAGGCAGGAGTG 241

Db    256 cgatgagtc--ccc-gtcc-c--cttctgaacagcaggttgcccaggacacagaggaggt 309
          ||  ||| |  ||| | || |  |||||||  ||||||| ||||||||||||||||||||
Qy    242 CGGAGAGCCTGCCCTGCCCTCTGCTTCTGAGGAGCAGGTAGCCCAGGACACAGAGGAGGT 301

Db    310 ctttcgaagctacgttttttacctccaccagcaggaacaggagacccagggg 361
          || || |||||||||||||| ||| |||||||||||||| |  | |||
Qy    302 TTTCCGCAGCTACGTTTTTTACCGCCATCAGCAGGAACAGGAGGCTGAAGGG 353


RESULT   10
LOCUS       N83998        232 bp    mRNA            EST       01-APR-1996
DEFINITION  KK4271F Homo sapiens cDNA clone KK4271 5' similar to BAK FOR BCL-2
            HOMOLOGUE.
ACCESSION   N83998
NID         g1259623
KEYWORDS    EST.
SOURCE      human clone=KK4271 primer=GAAATTAACCCTCACTAAAGGG library=Fetal
            heart, Lambda ZAP Express vector=Lambda ZAP Express host=E. coli
            XL1-Blue Rsite1=EcoRI Rsite2=XhoI mRNA was purified from human
            fetal hearts (8-10 weeks).  cDNA was synthesized using a XhoI-Oligo
            dT adaptor-primer.  EcoRI adaptors were ligated, followed by
            digestion with XhoI, for directional cloning into predigested
            lambda ZAP Express.
  ORGANISM  Homo sapiens
            Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata;
            Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
            Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria;
            Eutheria; Archonta; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 232)
  AUTHORS   Liew,C.C.
  TITLE     cDNAs from fetal heart
  JOURNAL   Unpublished (1996)
COMMENT
            Contact: Liew CC
            Molecular Cardiology
            University of Toronto
            Banting Institute, 100 College St., Toronto, Ontario, M5G1L5
            Tel: 4169788758
            Fax: 4169785650
            Email: liewcc@utcc.utoronto.ca.
FEATURES             Location/Qualifiers
     source          1..232
                     /organism="Homo sapiens"
                     /clone="KK4271"
     mRNA            <1..>232
BASE COUNT       52 a     69 c     57 g     54 t
ORIGIN

  Query Match            3.0%;  Score 62;  DB 133;  Length 232;
  Best Local Similarity  98.4%;  Pred. No. 4.39e-67;
  Matches    63;  Conservative    0;  Mismatches   1;  Indels   0;  Gaps   0;

Db     87 tcagaggagcaggtagcccaggacacagaggaggttttccgcagctacgttttttaccgc 146
          || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    267 TCTGAGGAGCAGGTAGCCCAGGACACAGAGGAGGTTTTCCGCAGCTACGTTTTTTACCGC 326

Db    147 catc 150
          ||||
Qy    327 CATC 330


RESULT   11
LOCUS       W82028        270 bp    mRNA            EST       12-SEP-1996
DEFINITION  mf02e01.r1 Soares mouse p3NMF19.5 Mus musculus cDNA clone 403896
            5'.
ACCESSION   W82028
NID         g1393105
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Mus.
REFERENCE   1  (bases 1 to 270)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT
            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:247664
            Seq primer: -28M13 rev2 from Amersham.
FEATURES             Location/Qualifiers
     source          1..270
                     /organism="Mus musculus"
                     /note="Vector: pT7T3D (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCATTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was size selected, ligated to Eco RI
                     adapters (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of a modified pT7T3 vector
                     (Pharmacia). Library went through one round of
                     normalization to a Cot = 5. Library constructed by Bento
                     Soares and M.Fatima Bonaldo. RNA was kindly provided by
                     Dr. Minoru Ko (Wayne State University)."
                     /clone="403896"
                     /clone_lib="Soares mouse p3NMF19.5"
                     /dev_stage="19.5 dpc total fetus"
                     /lab_host="DH10B (ampicillin resistant)"
     mRNA            <1..>270
BASE COUNT       61 a     83 c     74 g     52 t
ORIGIN

  Query Match            2.2%;  Score 46;  DB 175;  Length 270;
  Best Local Similarity  76.7%;  Pred. No. 3.69e-38;
  Matches   102;  Conservative    0;  Mismatches  26;  Indels   5;  Gaps   4;

Db    143 ctccaccaagacctgaaaaatggcatctggacaaggaccaggtcccccgaaggtgg-gt- 200
          ||||||  ||||||||||||||| || || ||||| |||||||| || | |  || ||
Qy    182 CTCCACTGAGACCTGAAAAATGGCTTCGGGGCAAGGCCCAGGTCCTCCCAGGCAGGAGTG 241

Db    201 ctgcgatgagtccccgtcc-c--cttctgaacagcaggttgcccaggacacagaggaggt 257
          | | ||     ||| | || |  |||||||  ||||||| ||||||||||||||||||||
Qy    242 CGGAGAGCCTGCCCTGCCCTCTGCTTCTGAGGAGCAGGTAGCCCAGGACACAGAGGAGGT 301

Db    258 ctttcgaagctac 270
          || || |||||||
Qy    302 TTTCCGCAGCTAC 314


RESULT   12
LOCUS       W77319        445 bp    mRNA            EST       20-JUN-1996
DEFINITION  me64c04.r1 Soares mouse embryo NbME13.5 14.5 Mus musculus cDNA
            clone 400326 5'.
ACCESSION   W77319
NID         g1387344
KEYWORDS    EST.
SOURCE      house mouse.
```

```
  ORGANISM  Mus musculus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae;
            Murinae; Mus.
REFERENCE   1  (bases 1 to 445)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT
            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:244094
            Seq primer: ETPrimer
            High quality sequence stop: 341.
FEATURES             Location/Qualifiers
     source          1..445
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGGAAATTTTTTTTTTTTTTTTTTTTTT
                     T 3'], on equal amounts of mRNA from 2 13.5dpc and 2
                     14.5dpc embryos [total RNA provided by  Minoru Ko, Wayne
                     State Univ., from 2 ]; double-stranded cDNA was ligated to
                     Eco RI adaptors (Pharmacia), digested with Not I and
                     cloned into the Not I and Eco RI sites of the modified
                     pT7T3 vector. Library went through one round of
                     normalization, and was constructed by Bento Soares and
                     M.Fatima Bonaldo."
                     /clone="400326"
                     /clone_lib="Soares mouse embryo NbME13.5 14.5"
                     /sex="unknown"
                     /tissue_type="embryo"
                     /dev_stage="13.5-14.5dpc total fetus"
                     /lab_host="DH10B"
     mRNA            <1..>445
BASE COUNT       91 a    155 c     95 g    104 t
ORIGIN

  Query Match             2.1%;  Score 44;  DB 126;  Length 445;
  Best Local Similarity  72.1%;  Pred. No. 9.87e-35;
  Matches    80;  Conservative    0;  Mismatches  30;  Indels   1;  Gaps   1;

Db     266 tgccttggcggctggactctcagggattctgagcccagagtgagggcagaggtgagagtt 325
           |||||| || | |||||||||||||||||||| ||   | ||| |||  | ||||
Qy    1441 TGCCTTTGCAGTTGGACTCTCAGGGATTCTGGGCTTGGGGTGTGGGGTGGGGTGGAGTCG 1500

Db     326 cagaccacagctgtctgaacacatccatcaggctc-tccaatcctgtgtcc 375
           ||||||| |||||||||||| ||    ||||   | ||||| |||| |||
Qy    1501 CAGACCAGAGCTGTCTGAACTCACGTGTCAGAAGCCTCCAAGCCTGCCTCC 1551


RESULT   13
LOCUS       AA139013      320 bp    mRNA            EST       02-DEC-1996
DEFINITION  mr04a06.r1 Soares mouse 3NbMS Mus musculus cDNA clone 596434 5'
            similar to TR:G595926 G595926 BAK-2 PROTEIN. ;.
ACCESSION   AA139013
NID         g1701240
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Mus.
REFERENCE   1  (bases 1 to 320)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT
            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:361866
            Seq primer: -28M13 rev2 from Amersham.
FEATURES             Location/Qualifiers
     source          1..320
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGCTGTTTTTTTTTTTTTTTTTTTTTT
                     3']; double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not
                     I and Eco RI sites of the modified pT7T3 vector. RNA
                     provided by Dr. Bertrand Jordan. Library went through
                     three rounds of normalization, and was constructed by
                     Bento Soares and M.Fatima Bonaldo."
                     /clone="596434"
                     /clone_lib="Soares mouse 3NbMS"
                     /sex="male"
                     /dev_stage="4 weeks"
                     /lab_host="DH10B"
     mRNA            <1..>320
BASE COUNT       68 a     92 c     96 g     64 t
ORIGIN

  Query Match             1.9%;  Score 39;  DB 186;  Length 320;
  Best Local Similarity  84.2%;  Pred. No. 2.04e-26;
  Matches    48;  Conservative    0;  Mismatches   9;  Indels   0;  Gaps   0;

Db     264 ccgtccccttctgaacagcaggttgcccaggacacagaggaggtctttcgaagctac 320
           || ||  |||||||  ||||||| |||||||||||||||||||| || || ||||||
Qy     258 CCCTCTGCTTCTGAGGAGCAGGTAGCCCAGGACACAGAGGAGGTTTTCCGCAGCTAC 314


RESULT   14
LOCUS       H31839        408 bp    mRNA            EST       08-SEP-1995
DEFINITION  EST106322 Rattus sp. cDNA 3' end.
ACCESSION   H31839
NID         g977256
KEYWORDS    EST.
SOURCE      rat primer=M13 - 21 library=Rat PC-12 cells, untreated
            vector=pBluescript SK- Rsite1=EcoRI Rsite2=XhoI poly(A)+ RNA was
            purified from untreated PC12 cells cultured for 9 days.  cDNA was
            constructed using an oligo-dT primer and directionally cloned using
            the Lambda ZAP II Vector Kit by Stratagene.
  ORGANISM  Rattus sp.
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Myomorpha;
            Muridae; Murinae; Rattus.
REFERENCE   1  (bases 1 to 408)
  AUTHORS   Lee,N.H., Weinstock,K.G., Kirkness,E.F., Earle-Hughes,J.A.,
            Fuldner,R.A., Marmaras,S., Glodek,A., Gocayne,J.D., Adams,M.D.,
            Kerlavage,A.R., Fraser,C.M. and Venter,J.C.
```

```
  TITLE     Comparative expressed sequence tag analysis of differential gene
            expression profiles in PC-12 cells before and after nerve growth
            factor treatment
  JOURNAL   Proc. Natl. Acad. Sci. USA 92, 8303-8307 (1995)
COMMENT     Other_ESTs: EST106323
            Contact: Lee NH
            The Institute for Genomic Research
            932 Clopper Rd, Gaithersburg, MD 20878
            Tel: 3018699056
            Fax: 3018699423
            Email: nhlee@tigr.org
            For clone availability please contact  the TIGR Database
            (tdbinfo@tdb.tigr.org).
FEATURES             Location/Qualifiers
     source          1..408
                     /organism="Rattus sp."
     mRNA            <1..>408
BASE COUNT       89 a    113 c    115 g     90 t      1 others
ORIGIN

  Query Match           1.6%;  Score 33;  DB 82;  Length 408;
  Best Local Similarity  71.4%;  Pred. No. 5.04e-17;
  Matches    80;  Conservative    0;  Mismatches  29;  Indels   3;  Gaps   3;

Db      7 ttgggggcattgcaccgtttatttccagacaccgag-ggcgaggtgggggaatccat-ct 64
          ||| ||| ||||||| |||||||||||| ||||  || ||  ||| || ||  |     ||
Cp   2077 TTGAGGGGATTGCACAGTTTATTTCCAAACACTCAGAGGATAGGGGGTGGCCTATGGGCT 2018

Db     65 gcctccccccaccttgctacatttggcttaatcaagacttttctcacccct 116
            | ||  ||  ||||  || |||||||||| ||||||||  || ||| |||||
Cp   2017 CCATCTGCCTCCCCTCCCTGCATTTGGCTGAATCAAGAACTT-CTCCCCCCT 1967


RESULT   15
LOCUS       H31840        279 bp    mRNA            EST       08-SEP-1995
DEFINITION  EST106323 Rattus sp. cDNA 5' end.
ACCESSION   H31840
NID         g977257
KEYWORDS    EST.
SOURCE      rat primer=M13 Reverse library=Rat PC-12 cells, untreated
            vector=pBluescript SK- Rsitel=EcoRI Rsite2=XhoI poly(A)+ RNA was
            purified from untreated PC12 cells cultured for 9 days.  cDNA was
            constructed using an oligo-dT primer and directionally cloned using
            the Lambda ZAP II Vector Kit by Stratagene.
  ORGANISM  Rattus sp.
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Myomorpha;
            Muridae; Murinae; Rattus.
REFERENCE   1  (bases 1 to 279)
  AUTHORS   Lee,N.H., Weinstock,K.G., Kirkness,E.F., Earle-Hughes,J.A.,
            Fuldner,R.A., Marmaras,S., Glodek,A., Gocayne,J.D., Adams,M.D.,
            Kerlavage,A.R., Fraser,C.M. and Venter,J.C.
  TITLE     Comparative expressed sequence tag analysis of differential gene
            expression profiles in PC-12 cells before and after nerve growth
            factor treatment
  JOURNAL   Proc. Natl. Acad. Sci. USA 92, 8303-8307 (1995)
COMMENT     Other_ESTs: EST106322
            Contact: Lee NH
            The Institute for Genomic Research
            932 Clopper Rd, Gaithersburg, MD 20878
            Tel: 3018699056
            Fax: 3018699423
            Email: nhlee@tigr.org
            For clone availability please contact  the TIGR Database
            (tdbinfo@tdb.tigr.org).
FEATURES             Location/Qualifiers
     source          1..279
                     /organism="Rattus sp."
     mRNA            <1..>279
BASE COUNT       65 a     80 c     76 g     53 t      5 others
ORIGIN

  Query Match           1.5%;  Score 31;  DB 82;  Length 279;
  Best Local Similarity  82.7%;  Pred. No. 4.50e-14;
  Matches    43;  Conservative    0;  Mismatches   8;  Indels   1;  Gaps   1;

Db    226 cctccaccaagaccggaaaaatggcatccgggacaaggaccaggtnctncca 277
          |||||||  ||||| |||||||||| || ||| ||||| |||||| || |||
Qy    181 CCTCCACTGAGACCTGAAAAATGGCTTCGGGG-CAAGGCCCAGGTCCTCCCA 231


Search completed: Thu Aug 21 11:29:33 1997
Job time : 3043 secs.
```

****



****

Release 2.1D John F. Collins, Biocomputing Research Unit.

MPsrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Thu Aug 21 11:29:58 1997; MasPar time 554.19 Seconds
1038.366 Million cell updates/sec

Tabular output not generated.

Title: >US-08-320-157-6
Description: (1-2094) from US08320157.seq
Perfect Score: 2094
N.A. Sequence: 1 GAGGATCTACAGGGGACAAG.....CAAAAAAAAAACGGAGATCC 2094
Comp: CTCCTAGATGTCCCCTGTTC.....GTTTTTTTTTTGCCTCTAGG

Scoring table: TABLE default
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 359085 seqs, 137405154 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: EST-STS-THREE
1:EST199 2:EST200 3:EST201 4:EST202 5:EST203 6:EST204
7:EST205 8:EST206 9:EST207 10:EST208 11:EST209 12:EST210
13:EST211 14:EST212 15:EST213 16:EST214 17:EST215
18:EST216 19:EST217 20:EST218 21:EST219 22:EST220
23:EST221 24:EST222 25:EST223 26:EST224 27:EST225
28:EST226 29:EST227 30:EST228 31:EST229 32:EST230
33:EST231 34:EST232 35:EST233 36:STS1 37:STS2 38:STS3
39:STS4 40:STS5 41:STS6 42:STS7 43:STS8 44:STS9 45:STS10
46:STS11 47:STS12 48:STS13

Database: EST-STS-FOUR
49:gnEST1 50:gnEST2 51:gnEST3 52:gnEST4 53:gnEST5
54:gnEST6 55:gnEST7 56:gnEST8 57:gnEST9 58:gnEST10
59:gnEST11 60:gnEST12 61:gnEST13 62:gnEST14 63:gnEST15
64:gnEST16 65:gnEST17 66:gnEST18 67:gnEST19 68:gnEST20
69:gnEST21 70:gnEST22 71:gnEST23 72:gnEST24 73:gnEST25
74:gnSTS 75:enEST1 76:enEST2 77:enEST3 78:enEST4
79:enEST5 80:enEST6 81:enEST7 82:enEST8 83:enEST9
84:enEST10 85:enEST11 86:enEST12 87:enEST13 88:enEST14
89:enEST15 90:enEST16 91:enEST17 92:enEST18 93:enEST19
94:enEST20 95:enEST21 96:enEST22 97:enEST23 98:enEST24
99:enEST25 100:enEST26 101:enEST27 102:enEST28
103:enEST29 104:enEST30 105:enEST31 106:enSTS 107:ueEST1
108:ueEST2

Statistics: Mean 11.796; Variance 2.211; scale 5.335

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	% Query	No.	Score	Match	Length	DB	ID	Description	Pred. No.
		1	1248	59.6	1258	38	G06794	human STS WI-7983.	0.00e+00
		2	154	7.4	161	88	HS886334	zb46a10.r1 Soares fet	1.68e-256
		3	55	2.6	133	91	MM1142620	mt62a08.r1 Soares 2Nb	9.33e-55
		4	55	2.6	133	61	AA185611	mt62a08.r1 Soares 2Nb	9.33e-55
		5	55	2.6	133	13	AA185611	mt62a08.r1 Soares 2Nb	9.33e-55
		6	44	2.1	309	60	AA222196	mw21g06.r1 Soares mou	3.05e-35
		7	44	2.1	309	91	MM1141069	mw21g06.r1 Soares mou	3.05e-35
		8	42	2.0	498	67	AA242472	mx26g11.r1 Soares mou	7.47e-32
		9	42	2.0	498	92	MM1158998	mx26g11.r1 Soares mou	7.47e-32
		10	40	1.9	310	7	AA123834	mp93c04.r1 Soares 2Nb	1.60e-28
		11	40	1.9	310	98	MMAA23834	mp93c04.r1 Soares 2Nb	1.60e-28
		12	40	1.9	310	58	AA123834	mp93c04.r1 Soares 2Nb	1.60e-28
		13	39	1.9	320	56	AA139013	mr04a06.r1 Soares mou	6.98e-27
		14	39	1.9	331	108	MM1139075	my25f11.r1 Barstead m	6.98e-27
		15	39	1.9	331	49	AA220617	my25f11.r1 Barstead m	6.98e-27
		16	39	1.9	341	91	MM1141049	mw21a06.r1 Soares mou	6.98e-27
		17	39	1.9	341	60	AA222225	mw21a06.r1 Soares mou	6.98e-27
c		18	25	1.2	410	8	AA129872	zl14c11.s1 Soares pre	2.71e-06
c		19	23	1.1	99	25	N58693	yv64f10.r1 Soares fet	7.19e-04
c		20	23	1.1	156	36	DM36D7S	D. melanogaster STS d	7.19e-04
		21	22	1.1	222	53	T02955	FB17G10 Fetal brain,	9.94e-03
c		22	22	1.1	250	41	G15328	human STS SHGC-16013.	9.94e-03
		23	22	1.1	347	55	AA089066	mo63f04.r1 Stratagene	9.94e-03
c		24	24	1.1	412	25	N59408	yv51e02.s1 Soares fet	4.65e-05
		25	22	1.1	418	33	AA110168	mp42f02.r1 Barstead M	9.94e-03
		26	22	1.1	424	89	HSW3043	ze15d02.s1 Soares fet	9.94e-03
		27	22	1.1	424	1	W92304	ze15d02.s1 Soares fet	9.94e-03
		28	22	1.1	425	25	N58738	yv74d09.s1 Soares fet	9.94e-03
c		29	24	1.1	442	8	AA126907	zl87b04.s1 Stratagene	4.65e-05
		30	22	1.1	450	25	N57654	yv54d04.s1 Soares fet	9.94e-03
c		31	24	1.1	454	29	AA033526	zf01b10.s1 Soares fet	4.65e-05
c		32	22	1.1	460	21	AA018827	ze57e11.s1 Soares ret	9.94e-03
c		33	23	1.1	465	42	G21837	human STS WI-11990.	7.19e-04
c		34	22	1.1	467	21	AA020955	ze64h11.s1 Soares ret	9.94e-03
		35	23	1.1	501	86	HS1174584	nc34h03.r1 NCI CGAP P	7.19e-04
		36	23	1.1	501	73	AA259189	nc34h03.r1 NCI CGAP P	7.19e-04
c		37	22	1.1	534	104	MMAA85558	mt62e08.r1 Soares 2Nb	9.94e-03
c		38	22	1.1	534	61	AA185558	mt62e08.r1 Soares 2Nb	9.94e-03
c		39	22	1.1	534	13	AA185558	mt62e08.r1 Soares 2Nb	9.94e-03
		40	24	1.1	553	4	AA115608	zk89a06.r1 Soares pre	4.65e-05
c		41	24	1.1	560	4	AA115474	zk89a06.s1 Soares pre	4.65e-05
		42	22	1.1	600	25	N58258	yv67g01.s1 Soares fet	9.94e-03
c		43	22	1.1	610	41	G15924	human STS CHLC.UTR_03	9.94e-03
		44	22	1.1	676	25	N53559	yv43f09.s1 Soares fet	9.94e-03
c		45	22	1.1	1124	38	G07120	human STS WI-9107.	9.94e-03

ALIGNMENTS

RESULT 1
LOCUS G06794 1258 bp DNA STS 19-OCT-1995
DEFINITION human STS WI-7983.
ACCESSION G06794
NID g860039
KEYWORDS STS sequence; primer; sequence tagged site.
SOURCE human STSs derived from sequences in dbEST and the Unigene collection.
ORGANISM Homo sapiens
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Archonta; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1258)
AUTHORS Hudson,T.
TITLE Whitehead Institute/MIT Center for Genome Research; Physically Mapped ESTs
JOURNAL Unpublished (1995)
COMMENT
Contact: Thomas Hudson
Whitehead Institute/MIT Center for Genome Research
Whitehead Institute for Biomedical Research

9 Cambridge Center, Cambridge MA 02142 USA
Tel: 617 252 1900
Fax: 617 252 1902
Email: thudson@genome.wi.mit.edu

Primer A: CTGATAACTTGGGGAGGCAA
Primer B: GAGAGTCCAACTGCAAAGGC
STS size: 345
PCR Profile:
Presoak:
Denaturation:
Annealing: 56 degrees C
Polymerization:
PCR Cycles: 35
Thermal Cycler:
Protocol:
Template: 10 ng
Primer: each 5 pM
dNTPs: each 4 nM
Taq Polymerase: 0.025 units/ul
Total Vol: 20 ul

Buffer:
MgCl2: 1.5 mM
KCl: 50 mM
Tris-HCL: 10 mM
pH: 9.3

Prepared with primer pairs derived from U16811 -- Unigene.

```
FEATURES             Location/Qualifiers
     source          1..1258
                     /organism="Homo sapiens"
     STS             281..625
                     /map="751_A_7; 787_E_4; 908_C_6; 923_D_2; 919_D_10;
                     942_A_12; 808_G_(6,11)"
     primer_bind     281..300
                     /map="751_A_7; 787_E_4; 908_C_6; 923_D_2; 919_D_10;
                     942_A_12; 808_G_(6,11)"
     primer_bind     complement(606..625)
                     /map="751_A_7; 787_E_4; 908_C_6; 923_D_2; 919_D_10;
                     942_A_12; 808_G_(6,11)"
BASE COUNT      236 a    356 c    355 g    301 t     10 others
ORIGIN
```

Query Match 59.6%; Score 1248; DB 38; Length 1258;
Best Local Similarity 99.2%; Pred. No. 0.00e+00;
Matches 1248; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

```
Db      1 ctcccaagggtgccctttgggtcccggttcagacccctgcctggacttaagcgaagtctt 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    837 CTCCCAAGGGTGCCCTTTGGGTCCCGGTTCAGACCCCTGCCTGGACTTAAGCGAAGTCTT 896

Db     61 tgccttctctgttcccttgcagggtccccccctcaagagtacagaagctttagcaagtgtg 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    897 TGCCTTCTCTGTTCCCTTGCAGGGTCCCCCCTCAAGAGTACAGAAGCTTTAGCAAGTGTG 956

Db    121 cactccagcttcggaggccctgcgtgggggccagtcaggctgcagaggcacctcaacatt 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    957 CACTCCAGCTTCGGAGGCCCTGCGTGGGGGCCAGTCAGGCTGCAGAGGCACCTCAACATT 1016

Db    181 gcatggtgctagtgccctctctctgggcccagggctgtggccgtctcctccctcagctct 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1017 GCATGGTGCTAGTGCCCTCTCTCTGGGCCCAGGGCTGTGGCCGTCTCCTCCCTCAGCTCT 1076

Db    241 ctgggacctccttagccctgtctgctaggcgctggggagactgataacttggggaggcaa 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1077 CTGGGACCTCCTTAGCCCTGTCTGCTAGGCGCTGGGGAGACTGATAACTTGGGGAGGCAA 1136

Db    301 gagactgggagccacttctccccagaaagtgtttaacggttttagctttttataataccc 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1137 GAGACTGGGAGCCACTTCTCCCCAGAAAGTGTTTAACGGTTTTAGCTTTTTATAATACCC 1196

Db    361 ttgtgagagcccattcccaccattctacctgaggccaggacgtctggggtgtggggattg 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1197 TTGTGAGAGCCCATTCCCACCATTCTACCTGAGGCCAGGACGTCTGGGGTGTGGGGATTG 1256

Db    421 gtgggtctatgttccccaggattcagctattctggaagatcagcaccctaagagatggga 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1257 GTGGGTCTATGTTCCCCAGGATTCAGCTATTCTGGAAGATCAGCACCCTAAGAGATGGGA 1316

Db    481 ctaggacctgagcctggtcctggccgtccctaagcatgtgtcccaggagcaggacctact 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1317 CTAGGACCTGAGCCTGGTCCTGGCCGTCCCTAAGCATGTGTCCCAGGAGCAGGACCTACT 1376

Db    541 aggagaggggggccaaggtcctgctcaactctacccctgctcccattcctccctccggcc 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1377 AGGAGAGGGGGGCCAAGGTCCTGCTCAACTCTACCCCTGCTCCCATTCCTCCCTCCGGCC 1436

Db    601 atactgcctttgcagttggactctcagggattctgggcttggggtgtggggtggggtgga 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1437 ATACTGCCTTTGCAGTTGGACTCTCAGGGATTCTGGGCTTGGGGTGTGGGGTGGGGTGGA 1496

Db    661 gtcgcagaccagagctgtctgaactcacgtgtcagaagcctccaagcctgcctcccaagg 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1497 GTCGCAGACCAGAGCTGTCTGAACTCACGTGTCAGAAGCCTCCAAGCCTGCCTCCCAAGG 1556

Db    721 tcctctcagttctctcccttcctctctccttatagacacttgctcccaacccattcacta 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1557 TCCTCTCAGTTCTCTCCCTTCCTCTCTCCTTATAGACACTTGCTCCCAACCCATTCACTA 1616

Db    781 caggtgaaggctctcacccatccctgggggccttgggtgagtggcctgctaaggctcctc 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1617 CAGGTGAAGGCTCTCACCCATCCCTGGGGGCCTTGGGTGAGTGGCCTGCTAAGGCTCCTC 1676

Db    841 cttgcccagactacagggcttaggacttggtttgttatatcagggaaaaggagtagggag 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1677 CTTGCCCAGACTACAGGGCTTAGGACTTGGTTTGTTATATCAGGGAAAAGGAGTAGGGAG 1736

Db    901 ttcatctggagggttctaagtgggagaaggactatcaacaccactaggaatcccagaggt 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1737 TTCATCTGGAGGGTTCTAAGTGGGAGAAGGACTATCAACACCACTAGGAATCCCAGAGGT 1796

Db    961 ggatcctccctcatggctctggcacagtgtaatccaggggtgtagatgggggaactgtga 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1797 GGATCCTCCCTCATGGCTCTGGCACAGTGTAATCCAGGGGTGTAGATGGGGGAACTGTGA 1856

Db   1021 atacttgaactctgttcccccacccctccatgctcctcacctgtctaggtctcctcagggt 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1857 ATACTTGAACTCTGTTCCCCCACCCTCCATGCTCCTCACCTGTCTAGGTCTCCTCAGGGT 1916

Db   1081 gggggggtgacagtgccttctctattggcacagcctagggtcttgggggtcaggggggaga 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1917 GGGGGGTGACAGTGCCTTCTCTATTGGCACAGCCTAGGGTCTTGGGGGTCAGGGGGGAGA 1976

Db   1141 agttcttgattcagccaaatgcagggaggggaggcagatggagcccataggccacccccct 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1977 AGTTCTTGATTCAGCCAAATGCAGGGAGGGGAGGCAGATGGAGCCCATAGGCCACCCCCT 2036

Db   1201 atcctctgagtgtttggaaataaactgtgcaatcccctcnnnnnnnnnnnnnggagatcc 1258
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   2037 ATCCTCTGAGTGTTTGGAAATAAACTGTGCAATCCCCTCAAAAAAAAAACGGAGATCC 2094
```

```
RESULT    2
ID   HS886334   standard; RNA; EST; 161 BP.
AC   W23886;
NI   g1300701
DT   09-MAY-1996 (Rel. 47, Created)
DT   07-MAR-1997 (Rel. 51, Last updated, Version 2)
DE   zb46a10.r1 Soares fetal lung NbHL19W Homo sapiens cDNA clone 306618
DE   5'.
KW   EST.
OS   Homo sapiens (human)
OC   Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
OC   Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
```

```
RP   1-161
RA   Hillier L., Clark N., Dubuque T., Elliston K., Hawkins M.,
RA   Holman M., Hultman M., Kucaba T., Le M., Lennon G., Marra M.,
RA   Parsons J., Rifkin L., Rohlfing T., Tan F., Trevaskis E.,
RA   Waterston R., Williamson A., Wohldmann P., Wilson R.;
RT   "WashU-Merck EST Project";
RL   Unpublished.
CC   Contact: Wilson RK WashU-Merck EST Project Washington University
CC   School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis,
CC   MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email:
CC   est@watson.wustl.edu This clone is available royalty-free through
CC   LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for
CC   further information. Insert Length: 988 Std Error: 0.00 Seq primer:
CC   mob.REGA+ET.
FH   Key             Location/Qualifiers
FH
FT   source          1..161
FT                   /organism="Homo sapiens"
FT                   /note="Organ: lung; Vector: pT7T3D (Pharmacia) with a
FT                   modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
FT                   strand cDNA was primed with a Not I - oligo(dT) primer
FT                   [5'-TGTTACCAATCTGAAGTGGGAGCGGCCGCAATTTTTTTTTTTTTTTT-3'],
FT                   double-stranded cDNA was size selected, ligated to Eco RI
FT                   adapters (Pharmacia), digested with Not I and cloned into
FT                   the Not I and Eco RI sites of a modified pT7T3 vector
FT                   (Pharmacia). Library went through one round of
FT                   normalization to a Cot = 5. Library constructed by Bento
FT                   Soares and M.Fatima Bonaldo. This library was constructed
FT                   from the same fetus as the fetal heart library, Soares
FT                   fetal heart NbHH19W."
FT                   /clone="306618"
FT                   /clone_lib="Soares fetal lung NbHL19W"
FT                   /dev_stage="19 weeks"
FT                   /lab_host="DH10B (ampicillin resistant)"
FT   mRNA            <1..>161
SQ   Sequence 161 BP; 33 A; 46 C; 38 G; 44 T; 0 other;

  Query Match             7.4%;  Score 154;  DB 88;  Length 161;
  Best Local Similarity  99.4%;  Pred. No. 1.68e-256;
  Matches   160;  Conservative    0;  Mismatches   0;  Indels   1;  Gaps   1;

Db      1 cctcccctcagctctctgggacctccttagccctgtctgctaggcgctggggagactgata 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1063 CCTCCCTCAGCTCTCTGGGACCTCCTTAGCCCTGTCTGCTAGGCGCTGGGGAGACTGATA 1122

Db     61 acttggggaggcaagagactgggagccacttctccccagaaagtgtttaacggttttagc 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1123 ACTTGGGGAGGCAAGAGACTGGGAGCCACTTCTCCCCAGAAAGTGTTTAACGGTTTTAGC 1182

Db    121 tttttataataccccttgtgagaagcccattcccaccattct 161
          ||||||||||||||||||||| ||||||||||||||||||
Qy   1183 TTTTTATAATACCCTTGTGAGA-GCCCATTCCCACCATTCT 1222


RESULT    3
ID   MM1142620  standard; RNA; EST; 133 BP.
AC   AA185611;
NI   g1769266
DT   21-FEB-1997 (Rel. 51, Created)
DT   22-FEB-1997 (Rel. 51, Last updated, Version 2)
DE   mt62a08.r1 Soares 2NbMT Mus musculus cDNA clone 634454 5' similar
DE   to TR:G595926 G595926 BAK-2 PROTEIN. ;.
KW   EST.
OS   Mus musculus (house mouse)
OC   Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
OC   Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
OC   Mus.
RN   [1]
RP   1-133
RA   Marra M., Hillier L., Allen M., Bowles M., Dietrich N.,
RA   Dubuque T., Geisel S., Kucaba T., Lacy M., Le M., Martin J.,
RA   Morris M., Schellenberg K., Steptoe M., Tan F., Underwood K.,
RA   Moore B., Theising B., Wylie T., Lennon G., Soares B., Wilson R.,
RA   Waterston R.;
RT   "The WashU-HHMI Mouse EST Project";
RL   Unpublished.
CC   Contact: Marra M/Mouse EST Project WashU-HHMI Mouse EST Project
CC   Washington University School of MedicineP 4444 Forest Park Parkway,
CC   Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810
CC   Email: mouseest@watson.wustl.edu This clone is available
CC   royalty-free through LLNL ; contact the IMAGE Consortium
CC   (info@image.llnl.gov) for further information. MGI:386446 Trace
CC   considered overall poor quality Possible reversed clone: similarity
CC   on wrong strand Seq primer: -28M13 rev2 from Amersham High quality
CC   sequence stop: 1.
FH   Key             Location/Qualifiers
FH
FT   source          1..133
FT                   /organism="Mus musculus"
FT                   /strain="C57BL/6J"
FT                   /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
FT                   polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
FT                   was primed with a Not I - oligo(dT) primer [5'
FT                   TGTTACCAATCTGAAGTGGGAGCGGCCGCGTTTTTTTTTTTTTTTTTTTTTTTT
FT                   3']; double-stranded cDNA wasligated to Eco RI adaptors
FT                   (Pharmacia), digested with Not I and cloned into the Not I
FT                   and Eco RI sites of the modified pT7T3 vector. RNA provide
  d
FT                   by Dr. Bertrand Jordan. Library went through two rounds of
FT                   normalization, and was constructedby Bento Soares and
FT                   M.Fatima Bonaldo."
FT                   /clone="634454"
FT                   /clone_lib="Soares 2NbMT"
FT                   /sex="male"
FT                   /tissue_type="Thymus"
FT                   /dev_stage="4 weeks"
FT                   /lab_host="DH10B"
FT   mRNA            <1..>133
SQ   Sequence 133 BP; 33 A; 40 C; 34 G; 26 T; 0 other;

  Query Match             2.6%;  Score 55;  DB 91;  Length 133;
  Best Local Similarity  81.6%;  Pred. No. 9.33e-55;
  Matches    71;  Conservative    0;  Mismatches  16;  Indels   0;  Gaps   0;

Db     41 ccgtccccttctggacagcaggttgcccatgacacagaggaggtctttcgaagctacgta 100
          || ||  ||||||   ||||||| ||||| ||||||||||||||| || || ||||||||
Qy    258 CCCTCTGCTTCTGAGGAGCAGGTAGCCCAGGACACAGAGGAGGTTTTCCGCAGCTACGTT 317

Db    101 ttttaccttcaccagcaggaacatgag 127
          |||||||  || ||||||||||| |||
Qy    318 TTTTACCGCCATCAGCAGGAACAGGAG 344


RESULT    4
LOCUS       AA185611      133 bp    mRNA            EST       19-FEB-1997
DEFINITION  mt62a08.r1 Soares 2NbMT Mus musculus cDNA clone 634454 5' similar
            to TR:G595926 G595926 BAK-2 PROTEIN. ;.
ACCESSION   AA185611
NID         g1769266
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Mus.
REFERENCE   1  (bases 1 to 133)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT
            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
```

```
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:386446
            Trace considered overall poor quality
            Possible reversed clone: similarity on wrong strand
            Seq primer: -28M13 rev2 from Amersham
            High quality sequence stop: 1.
FEATURES              Location/Qualifiers
     source           1..133
                      /organism="Mus musculus"
                      /strain="C57BL/6J"
                      /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                      polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                      was primed with a Not I - oligo(dT) primer [5'
                      TGTTACCAATCTGAAGTGGGAGCGGCCGCGTTTTTTTTTTTTTTTTTTTTTTT
                      3']; double-stranded cDNA was ligated to Eco RI adaptors
                      (Pharmacia), digested with Not I and cloned into the Not I
                      and Eco RI sites of the modified pT7T3 vector. RNA
                      provided by Dr. Bertrand Jordan. Library went through two
                      rounds of normalization, and was constructed by Bento
                      Soares and M.Fatima Bonaldo."
                      /clone="634454"
                      /clone_lib="Soares 2NbMT"
                      /sex="male"
                      /tissue_type="Thymus"
                      /dev_stage="4 weeks"
                      /lab_host="DH10B"
     mRNA             <1..>133
BASE COUNT       33 a     40 c     34 g     26 t
ORIGIN

  Query Match             2.6%;  Score 55;  DB 61;  Length 133;
  Best Local Similarity  81.6%;  Pred. No. 9.33e-55;
  Matches    71;  Conservative    0;  Mismatches  16;  Indels   0;  Gaps   0;

Db     41 ccgtccccttctggacagcaggttgcccatgacacagaggaggtctttcgaagctacgta 100
          || ||  ||||||   ||||||| ||||| ||||||||||||| || || ||||||||
Qy    258 CCCTCTGCTTCTGAGGAGCAGGTAGCCCAGGACACAGAGGAGGTTTTCCGCAGCTACGTT 317

Db    101 ttttaccttcaccagcaggaacatgag 127
          |||||||  || ||||||||||| |||
Qy    318 TTTTACCGCCATCAGCAGGAACAGGAG 344


RESULT    5
LOCUS       AA185611      133 bp    mRNA            EST       07-JAN-1997
DEFINITION  mt62a08.r1 Soares 2NbMT Mus musculus cDNA clone 634454 5' similar
            to TR:G595926 G595926 BAK-2 PROTEIN. ;.
ACCESSION   AA185611
NID         g1769266
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Mus.
REFERENCE   1  (bases 1 to 133)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT
            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:386446
            Trace considered overall poor quality
            Possible reversed clone: similarity on wrong strand
            Seq primer: -28M13 rev2 from Amersham
            High quality sequence stop: 1.
FEATURES              Location/Qualifiers
     source           1..133
                      /organism="Mus musculus"
                      /strain="C57BL/6J"
                      /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                      polylinker; Site_1: Not I; Site_2: Eco RI; [5'
                      TGTTACCAATCTGAAGTGGGAGCGGCCGCGTTTTTTTTTTTTTTTTTTTTTTT
                      3']; double-stranded cDNA was ligated to Eco RI adaptors
                      (Pharmacia), digested with Not I and cloned into the Not
                      I and Eco RI sites of the modified pT7T3 vector. RNA
                      provided by Dr. Bertrand Jordan. Library went through two
                      rounds of normalization, and was constructed by Bento
                      Soares and M.Fatima Bonaldo."
                      /clone="634454"
                      /clone_lib="Soares 2NbMT"
                      /sex="male"
                      /dev_stage="4 weeks"
                      /lab_host="DH10B"
     mRNA             <1..>133
BASE COUNT       33 a     40 c     34 g     26 t
ORIGIN

  Query Match             2.6%;  Score 55;  DB 13;  Length 133;
  Best Local Similarity  81.6%;  Pred. No. 9.33e-55;
  Matches    71;  Conservative    0;  Mismatches  16;  Indels   0;  Gaps   0;

Db     41 ccgtccccttctggacagcaggttgcccatgacacagaggaggtctttcgaagctacgta 100
          || ||  ||||||   ||||||| ||||| ||||||||||||| || || ||||||||
Qy    258 CCCTCTGCTTCTGAGGAGCAGGTAGCCCAGGACACAGAGGAGGTTTTCCGCAGCTACGTT 317

Db    101 ttttaccttcaccagcaggaacatgag 127
          |||||||  || ||||||||||| |||
Qy    318 TTTTACCGCCATCAGCAGGAACAGGAG 344


RESULT    6
LOCUS       AA222196      309 bp    mRNA            EST       18-FEB-1997
DEFINITION  mw21g06.r1 Soares mouse 3NME12 5 Mus musculus cDNA clone 671386 5'.
ACCESSION   AA222196
NID         g1842522
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Mus.
REFERENCE   1  (bases 1 to 309)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT
            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
```

```
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:411090
            Seq primer: -28m13 rev2 ET from Amersham
            High quality sequence stop: 122.
FEATURES              Location/Qualifiers
     source           1..309
                      /organism="Mus musculus"
                      /strain="C57BL/6J"
                      /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                      polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                      was primed with a Not I - oligo(dT) primer [5'
                      TGTTACCAATCTGAAGTGGGAGCGGCCGCCTTATTTTTTTTTTTTTTTTT 3'],
                      on total  mouse RNA [provided by Minoru Ko, Wayne State
                      Univ.]; double-stranded cDNA was ligated to Eco RI
                      adaptors (Pharmacia), digested with Not I and cloned into
                      the Not I and Eco RI sites of the modified pT7T3 vector.
                      Library went through one round of normalization, and was
                      constructed by Bento Soares and M. Fatima Bonaldo."
                      /clone="671386"
                      /clone_lib="Soares mouse 3NME12 5"
                      /sex="unknown"
                      /dev_stage="12.5dpc total fetus"
                      /lab_host="DH10B"
     mRNA             <1..>309
BASE COUNT       59 a     92 c     91 g     67 t
ORIGIN

  Query Match             2.1%;  Score 44;  DB 60;  Length 309;
  Best Local Similarity  72.1%;  Pred. No. 3.05e-35;
  Matches    80;  Conservative    0;  Mismatches  30;  Indels   1;  Gaps   1;

Db      32 tgccttggcggctggactctcagggattctgagcccagagtgagggcagaggtgagagtt 91
           |||||| || | |||||||||||||||||||| ||   | ||| |||  | ||||
Qy    1441 TGCCTTTGCAGTTGGACTCTCAGGGATTCTGGGCTTGGGGTGTGGGGTGGGGTGGAGTCG 1500

Db      92 cagaccacagctgtctgaacacatccatcaggctc-tccaatcctgtgtcc 141
           ||||||| |||||||||||| ||    ||||   | ||||| |||| |||
Qy    1501 CAGACCAGAGCTGTCTGAACTCACGTGTCAGAAGCCTCCAAGCCTGCCTCC 1551


RESULT    7
ID   MM1141069  standard; RNA; EST; 309 BP.
AC   AA222196;
NI   g1842522
DT   20-FEB-1997 (Rel. 51, Created)
DT   20-FEB-1997 (Rel. 51, Last updated, Version 1)
DE   mw21g06.r1 Soares mouse 3NME12 5 Mus musculus cDNA clone 671386 5'.
KW   EST.
OS   Mus musculus (house mouse)
OC   Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
OC   Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
OC   Mus.
RN   [1]
RP   1-309
RA   Marra M., Hillier L., Allen M., Bowles M., Dietrich N.,
RA   Dubuque T., Geisel S., Kucaba T., Lacy M., Le M., Martin J.,
RA   Morris M., Schellenberg K., Steptoe M., Tan F., Underwood K.,
RA   Moore B., Theising B., Wylie T., Lennon G., Soares B., Wilson R.,
RA   Waterston R.;
RT   "The WashU-HHMI Mouse EST Project";
RL   Unpublished.
CC   Contact: Marra M/Mouse EST Project WashU-HHMI Mouse EST Project
CC   Washington University School of MedicineP 4444 Forest Park Parkway,
CC   Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810
CC   Email: mouseest@watson.wustl.edu This clone is available
CC   royalty-free through LLNL ; contact the IMAGE Consortium
CC   (info@image.llnl.gov) for further information. MGI:411090 Seq
CC   primer: -28m13 rev2 ET from Amersham High quality sequence stop:
CC   122.
FH   Key             Location/Qualifiers
FH
FT   source          1..309
FT                   /organism="Mus musculus"
FT                   /strain="C57BL/6J"
FT                   /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
FT                   polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
FT                   was primed with a Not I - oligo(dT) primer [5'
FT                   TGTTACCAATCTGAAGTGGGAGCGGCCGCCTTATTTTTTTTTTTTTTTT 3'], o
  n
FT                   total mouse RNA [providedby Minoru Ko, Wayne State Univ.];
FT                   double-stranded cDNA was ligated to Eco RI adaptors
FT                   (Pharmacia), digested with Not I and cloned into the Not I
FT                   and Eco RI sites of the modified pT7T3 vector. Library wen
  t
FT                   through one round of normalization, and was constructed by
FT                   Bento Soares and M. Fatima Bonaldo."
FT                   /clone="671386"
FT                   /clone_lib="Soares mouse 3NME12 5"
FT                   /sex="unknown"
FT                   /dev_stage="12.5dpc total fetus"
FT                   /lab_host="DH10B"
FT   mRNA            <1..>309
SQ   Sequence 309 BP; 59 A; 92 C; 91 G; 67 T; 0 other;

  Query Match             2.1%;  Score 44;  DB 91;  Length 309;
  Best Local Similarity  72.1%;  Pred. No. 3.05e-35;
  Matches    80;  Conservative    0;  Mismatches  30;  Indels   1;  Gaps   1;

Db      32 tgccttggcggctggactctcagggattctgagcccagagtgagggcagaggtgagagtt 91
           |||||| || | |||||||||||||||||||| ||   | ||| |||  | ||||
Qy    1441 TGCCTTTGCAGTTGGACTCTCAGGGATTCTGGGCTTGGGGTGTGGGGTGGGGTGGAGTCG 1500

Db      92 cagaccacagctgtctgaacacatccatcaggctc-tccaatcctgtgtcc 141
           ||||||| |||||||||||| ||    ||||   | ||||| |||| |||
Qy    1501 CAGACCAGAGCTGTCTGAACTCACGTGTCAGAAGCCTCCAAGCCTGCCTCC 1551


RESULT    8
LOCUS       AA242472      498 bp    mRNA            EST       07-MAR-1997
DEFINITION  mx26g11.r1 Soares mouse NML Mus musculus cDNA clone 681380 5'.
ACCESSION   AA242472
NID         g1873149
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Mus.
REFERENCE   1  (bases 1 to 498)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT
            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:421084
            Seq primer: -28m13 rev2 ET from Amersham
            High quality sequence stop: 477.
FEATURES              Location/Qualifiers
     source           1..498
                      /organism="Mus musculus"
                      /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                      polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                      was primed with a Not I - oligo(dT) primer [5'
```

```
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGAATCTTTTTTTTTTTTTTTTTT 3'];
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     constructed and normalized by Bento Soares and M.Fatima
                     Bonaldo."
                     /clone="681380"
                     /clone_lib="Soares mouse NML"
                     /tissue_type="Liver"
                     /lab_host="DH10B"
     mRNA            <1..>498
BASE COUNT       99 a    166 c    124 g    109 t
ORIGIN

  Query Match             2.0%;  Score 42;  DB 67;  Length 498;
  Best Local Similarity  67.4%;  Pred. No. 7.47e-32;
  Matches    91;  Conservative    0;  Mismatches  43;  Indels   1;  Gaps   1;

Db     198 tctcctaccccaatccatccatcatttttgccttggggctggactctcagggattctgagccc 257
           || | | || |  |||  |||| |  | |||  | ||||||||||||||||||| ||
Qy    1417 TCCCATTCCTCCCTCCGGCCATACTGCCTTTGCAGTTGGACTCTCAGGGATTCTGGGCTT 1476

Db     258 agagtgagggcagaggtgagagttcagaccacagctgtctgaacacatccatcaggctc- 316
           | ||| |||  | ||||      ||||||| |||||||||||| ||    ||||   |
Qy    1477 GGGGTGTGGGGTGGGGTGGAGTCGCAGACCAGAGCTGTCTGAACTCACGTGTCAGAAGCC 1536

Db     317 tccaatcctgtgtcc 331
           ||||| ||||  |||
Qy    1537 TCCAAGCCTGCCTCC 1551


RESULT    9
ID    MM1158998  standard; RNA; EST; 498 BP.
AC    AA242472;
NI    g1873149
DT    11-MAR-1997 (Rel. 51, Created)
DT    11-MAR-1997 (Rel. 51, Last updated, Version 1)
DE    mx26g11.r1 Soares mouse NML Mus musculus cDNA clone 681380 5'.
KW    EST.
OS    Mus musculus (house mouse)
OC    Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
OC    Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
OC    Mus.
RN    [1]
RP    1-498
RA    Marra M., Hillier L., Allen M., Bowles M., Dietrich N.,
RA    Dubuque T., Geisel S., Kucaba T., Lacy M., Le M., Martin J.,
RA    Morris M., Schellenberg K., Steptoe M., Tan F., Underwood K.,
RA    Moore B., Theising B., Wylie T., Lennon G., Soares B., Wilson R.,
RA    Waterston R.;
RT    "The WashU-HHMI Mouse EST Project";
RL    Unpublished.
CC    Contact: Marra M/Mouse EST Project WashU-HHMI Mouse EST Project
CC    Washington University School of MedicineP 4444 Forest Park Parkway,
CC    Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810
CC    Email: mouseest@watson.wustl.edu This clone is available
CC    royalty-free through LLNL ; contact the IMAGE Consortium
CC    (info@image.llnl.gov) for further information. MGI:421084 Seq
CC    primer: -28m13 rev2 ET from Amersham High quality sequence stop:
CC    477.
FH    Key             Location/Qualifiers
FH
FT    source          1..498
FT                    /organism="Mus musculus"
FT                    /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
FT                    polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
FT                    was primed with a Not I - oligo(dT) primer [5'
FT                    TGTTACCAATCTGAAGTGGGAGCGGCCGCGAATCTTTTTTTTTTTTTTTTT 3'];
FT                    double-stranded cDNA was ligated to Eco RI adaptors
FT                    (Pharmacia), digested with Not I and cloned into the Not I
FT                    and Eco RI sites of the modified pT7T3 vector. Library
FT                    constructed and normalized by Bento Soares and M.Fatima
FT                    Bonaldo."
FT                    /clone="681380"
FT                    /clone_lib="Soares mouse NML"
FT                    /tissue_type="Liver"
FT                    /lab_host="DH10B"
FT    mRNA            <1..>498
SQ    Sequence 498 BP; 99 A; 166 C; 124 G; 109 T; 0 other;

  Query Match             2.0%;  Score 42;  DB 92;  Length 498;
  Best Local Similarity  67.4%;  Pred. No. 7.47e-32;
  Matches    91;  Conservative    0;  Mismatches  43;  Indels   1;  Gaps   1;

Db     198 tctcctaccccaatccatccatcatttttgccttggggctggactctcagggattctgagccc 257
           || | | || |  |||  |||| |  | |||  | ||||||||||||||||||| ||
Qy    1417 TCCCATTCCTCCCTCCGGCCATACTGCCTTTGCAGTTGGACTCTCAGGGATTCTGGGCTT 1476

Db     258 agagtgagggcagaggtgagagttcagaccacagctgtctgaacacatccatcaggctc- 316
           | ||| |||  | ||||      ||||||| |||||||||||| ||    ||||   |
Qy    1477 GGGGTGTGGGGTGGGGTGGAGTCGCAGACCAGAGCTGTCTGAACTCACGTGTCAGAAGCC 1536

Db     317 tccaatcctgtgtcc 331
           ||||| ||||  |||
Qy    1537 TCCAAGCCTGCCTCC 1551


RESULT   10
LOCUS       AA123834      310 bp    mRNA            EST       21-NOV-1996
DEFINITION  mp93c04.r1 Soares 2NbMT Mus musculus cDNA clone 576774 5' similar
            to TR:G595926 G595926 BAK-2 PROTEIN. ;.
ACCESSION   AA123834
NID         g1682242
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Mus.
REFERENCE   1  (bases 1 to 310)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT
            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:351422
            Seq primer: -28M13 rev2 from Amersham.
FEATURES             Location/Qualifiers
     source          1..310
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGTTTTTTTTTTTTTTTTTTTTTTT
                     3']; double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not
                     I and Eco RI sites of the modified pT7T3 vector. RNA
                     provided by Dr. Bertrand Jordan. Library went through two
                     rounds of normalization, and was constructed by Bento
                     Soares and M.Fatima Bonaldo."
                     /clone="576774"
                     /clone_lib="Soares 2NbMT"
                     /sex="male"
                     /dev_stage="4 weeks"
```

```
                  /lab_host="DH10B"
     mRNA         <1..>310
BASE COUNT       72 a     96 c     86 g     55 t      1 others
ORIGIN

  Query Match             1.9%;  Score 40;  DB 7;  Length 310;
  Best Local Similarity  78.2%;  Pred. No. 1.60e-28;
  Matches   104;  Conservative    0;  Mismatches  22;  Indels   7;  Gaps   5;

Db    185 ctccaccaagacctgaaaaatggcatctggacaaggaccaggtcccccgaagtgggc-tg 243
          ||||||  |||||||||||||||| || || ||||| |||||||| || | |  ||  ||
Qy    182 CTCCACTGAGACCTGAAAAATGGCTTCGGGGCAAGGCCCAGGTCCTCCCAGGCAGGAGTG 241

Db    244 cgatgagtc--ccc-gtcc-c--cttctgaacagcaggttgcccaggacacagaggaggt 297
          ||  ||| |  ||| | || |  |||||||  ||||||| ||||||||||||||||||||
Qy    242 CGGAGAGCCTGCCCTGCCCTCTGCTTCTGAGGAGCAGGTAGCCCAGGACACAGAGGAGGT 301

Db    298 ctttcgaagctac 310
          || || ||||||
Qy    302 TTTCCGCAGCTAC 314


RESULT   11
ID   MMAA23834  standard; RNA; EST; 310 BP.
AC   AA123834;
NI   g1682242
DT   25-NOV-1996 (Rel. 50, Created)
DT   19-FEB-1997 (Rel. 51, Last updated, Version 2)
DE   mp93c04.r1 Soares 2NbMT Mus musculus cDNA clone 576774 5' similar
DE   to TR:G595926 G595926 BAK-2 PROTEIN. ;.
KW   EST.
OS   Mus musculus (house mouse)
OC   Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
OC   Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
OC   Mus.
RN   [1]
RP   1-310
RA   Marra M., Hillier L., Allen M., Bowles M., Dietrich N.,
RA   Dubuque T., Geisel S., Kucaba T., Lacy M., Le M., Martin J.,
RA   Morris M., Schellenberg K., Steptoe M., Tan F., Underwood K.,
RA   Moore B., Theising B., Wylie T., Lennon G., Soares B., Wilson R.,
RA   Waterston R.;
RT   "The WashU-HHMI Mouse EST Project";
RL   Unpublished.
CC   Contact: Marra M/Mouse EST Project WashU-HHMI Mouse EST Project
CC   Washington University School of MedicineP 4444 Forest Park Parkway,
CC   Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810
CC   Email: mouseest@watson.wustl.edu This clone is available
CC   royalty-free through LLNL ; contact the IMAGE Consortium
CC   (info@image.llnl.gov) for further information. MGI:351422 Seq
CC   primer: -28M13 rev2 from Amersham.
FH   Key             Location/Qualifiers
FH
FT   source          1..310
FT                   /organism="Mus musculus"
FT                   /strain="C57BL/6J"
FT                   /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
FT                   polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
FT                   was primed with a Not I - oligo(dT) primer [5'
FT                   TGTTACCAATCTGAAGTGGGAGCGGCCGCGTTTTTTTTTTTTTTTTTTTTTTTT
FT                   3']; double-stranded cDNA wasligated to Eco RI adaptors
FT                   (Pharmacia), digested with Not I and cloned into the Not I
FT                   and Eco RI sites of the modified pT7T3 vector. RNA provide
  d
FT                   by Dr. Bertrand Jordan. Library went through two rounds of
FT                   normalization, and was constructedby Bento Soares and
FT                   M.Fatima Bonaldo."
FT                   /clone="576774"
FT                   /clone_lib="Soares 2NbMT"
FT                   /sex="male"
FT                   /tissue_type="Thymus"
FT                   /dev_stage="4 weeks"
FT                   /lab_host="DH10B"
FT   mRNA            <1..>310
SQ   Sequence 310 BP; 72 A; 96 C; 86 G; 55 T; 1 other;

  Query Match             1.9%;  Score 40;  DB 98;  Length 310;
  Best Local Similarity  78.2%;  Pred. No. 1.60e-28;
  Matches   104;  Conservative    0;  Mismatches  22;  Indels   7;  Gaps   5;

Db    185 ctccaccaagacctgaaaaatggcatctggacaaggaccaggtcccccgaagtgggc-tg 243
          ||||||  |||||||||||||||| || || ||||| |||||||| || | |  ||  ||
Qy    182 CTCCACTGAGACCTGAAAAATGGCTTCGGGGCAAGGCCCAGGTCCTCCCAGGCAGGAGTG 241

Db    244 cgatgagtc--ccc-gtcc-c--cttctgaacagcaggttgcccaggacacagaggaggt 297
          ||  ||| |  ||| | || |  |||||||  ||||||| ||||||||||||||||||||
Qy    242 CGGAGAGCCTGCCCTGCCCTCTGCTTCTGAGGAGCAGGTAGCCCAGGACACAGAGGAGGT 301

Db    298 ctttcgaagctac 310
          || || ||||||
Qy    302 TTTCCGCAGCTAC 314


RESULT   12
LOCUS        AA123834      310 bp    mRNA            EST       17-FEB-1997
DEFINITION   mp93c04.r1 Soares 2NbMT Mus musculus cDNA clone 576774 5' similar
             to TR:G595926 G595926 BAK-2 PROTEIN. ;.
ACCESSION    AA123834
NID          g1682242
KEYWORDS     EST.
SOURCE       house mouse.
  ORGANISM   Mus musculus
             Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
             Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
             Mus.
REFERENCE    1  (bases 1 to 310)
  AUTHORS    Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
             Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
             Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
             Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
             Waterston,R.
  TITLE      The WashU-HHMI Mouse EST Project
  JOURNAL    Unpublished (1996)
COMMENT
             Contact: Marra M/Mouse EST Project
             WashU-HHMI Mouse EST Project
             Washington University School of MedicineP
             4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
             Tel: 314 286 1800
             Fax: 314 286 1810
             Email: mouseest@watson.wustl.edu
             This clone is available royalty-free through LLNL ; contact the
             IMAGE Consortium (info@image.llnl.gov) for further information.
             MGI:351422
             Seq primer: -28M13 rev2 from Amersham.
FEATURES             Location/Qualifiers
     source          1..310
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGTTTTTTTTTTTTTTTTTTTTTTTT
                     3']; double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. RNA
                     provided by Dr. Bertrand Jordan. Library went through two
                     rounds of normalization, and was constructed by Bento
                     Soares and M.Fatima Bonaldo."
                     /clone="576774"
                     /clone_lib="Soares 2NbMT"
                     /sex="male"
                     /tissue_type="Thymus"
                     /dev_stage="4 weeks"
                     /lab_host="DH10B"
     mRNA            <1..>310
```

```
BASE COUNT       72 a     96 c     86 g     55 t      1 others
ORIGIN

  Query Match             1.9%;  Score 40;  DB 58;  Length 310;
  Best Local Similarity  78.2%;  Pred. No. 1.60e-28;
  Matches   104;  Conservative    0;  Mismatches  22;  Indels   7;  Gaps   5;

Db    185 ctccaccaagacctgaaaaatggcatctggacaaggaccaggtcccccgaagtgggc-tg 243
          ||||||  ||||||||||||||||| || || ||||| |||||||| || | |  ||  ||
Qy    182 CTCCACTGAGACCTGAAAAATGGCTTCGGGGCAAGGCCCAGGTCCTCCCAGGCAGGAGTG 241

Db    244 cgatgagtc--ccc-gtcc-c--cttctgaacagcaggttgcccaggacacagaggaggt 297
          ||  ||| |  ||| | || |  |||||||  ||||||| ||||||||||||||||||
Qy    242 CGGAGAGCCTGCCCTGCCCTCTGCTTCTGAGGAGCAGGTAGCCCAGGACACAGAGGAGGT 301

Db    298 ctttcgaagctac 310
          || || ||||||
Qy    302 TTTCCGCAGCTAC 314


RESULT   13
LOCUS       AA139013      320 bp    mRNA            EST       16-FEB-1997
DEFINITION  mr04a06.r1 Soares mouse 3NbMS Mus musculus cDNA clone 596434 5'
            similar to TR:G595926 G595926 BAK-2 PROTEIN. ;.
ACCESSION   AA139013
NID         g1701240
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Mus.
REFERENCE   1  (bases 1 to 320)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT
            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:361866
            Seq primer: -28M13 rev2 from Amersham.
FEATURES             Location/Qualifiers
     source          1..320
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGCTGTTTTTTTTTTTTTTTTTTTTTTT
                     3']; double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. RNA
                     provided by Dr. Bertrand Jordan. Library went through
                     three rounds of normalization, and was constructed by
                     Bento Soares and M.Fatima Bonaldo."
                     /clone="596434"
                     /clone_lib="Soares mouse 3NbMS"
                     /sex="male"
                     /tissue_type="Spleen"
                     /dev_stage="4 weeks"
                     /lab_host="DH10B"
     mRNA            <1..>320
```

```
BASE COUNT       68 a     92 c     96 g     64 t
ORIGIN

  Query Match             1.9%;  Score 39;  DB 56;  Length 320;
  Best Local Similarity  84.2%;  Pred. No. 6.98e-27;
  Matches    48;  Conservative    0;  Mismatches   9;  Indels   0;  Gaps   0;

Db    264 ccgtccccttctgaacagcaggttgcccaggacacagaggaggtctttcgaagctac 320
          || ||  |||||||  ||||||| ||||||||||||||||||||| || || ||||||
Qy    258 CCCTCTGCTTCTGAGGAGCAGGTAGCCCAGGACACAGAGGAGGTTTTCCGCAGCTAC 314


RESULT   14
ID   MM1139075  standard; RNA; EST; 331 BP.
AC   AA220617;
NI   g1838404
DT   13-FEB-1997 (Rel. 50, Created)
DT   13-FEB-1997 (Rel. 50, Last updated, Version 1)
DE   my25f11.r1 Barstead mouse pooled organs MPLRB4 Mus musculus cDNA
DE   clone 696909 5'.
KW   EST.
OS   Mus musculus (house mouse)
OC   Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
OC   Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
OC   Mus.
RN   [1]
RP   1-331
RA   Marra M., Hillier L., Allen M., Bowles M., Dietrich N.,
RA   Dubuque T., Geisel S., Kucaba T., Lacy M., Le M., Martin J.,
RA   Morris M., Schellenberg K., Steptoe M., Tan F., Underwood K.,
RA   Moore B., Theising B., Wylie T., Lennon G., Soares B., Wilson R.,
RA   Waterston R.;
RT   "The WashU-HHMI Mouse EST Project";
RL   Unpublished.
CC   Contact: Marra M/Mouse EST Project WashU-HHMI Mouse EST Project
CC   Washington University School of MedicineP 4444 Forest Park Parkway,
CC   Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810
CC   Email: mouseest@watson.wustl.edu This clone is available
CC   royalty-free through LLNL ; contact the IMAGE Consortium
CC   (info@image.llnl.gov) for further information. MGI:430469 Seq
CC   primer: -28m13 rev2 ET from Amersham.
FH   Key             Location/Qualifiers
FH
FT   source          1..331
FT                   /organism="Mus musculus"
FT                   /strain="FVB/N"
FT                   /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
FT                   polylinker; Site_1: EcoRI; Site_2: NotI; 1st strand cDNA
FT                   was primed with a Not I - oligo(dT) primer [5'
FT                   TGTTACGAATCTGAAGTGGGAGCGGCCGCCCTTTTTTTTTTTTTTTTTTTTTTT
FT                   3']; double-stranded cDNA wasligated to Eco RI adaptors
FT                   [GTTGGATTCGGTACC], digested with Not I and cloned into the
FT                   Not I and Eco RI sites of the modified pT7T3 vector.
FT                   Library constructed by Bob Barstead."
FT                   /clone="696909"
FT                   /clone_lib="Barstead mouse pooled organs MPLRB4"
FT                   /sex="mixed"
FT                   /dev_stage="7 day"
FT                   /lab_host="DH10B"
FT   mRNA            <1..>331
SQ   Sequence 331 BP; 78 A; 104 C; 89 G; 60 T; 0 other;

  Query Match             1.9%;  Score 39;  DB 108;  Length 331;
  Best Local Similarity  84.2%;  Pred. No. 6.98e-27;
  Matches    48;  Conservative    0;  Mismatches   9;  Indels   0;  Gaps   0;

Db    275 ccgtccccttctgaacagcaggttgcccaggacacagaggaggtctttcgaagctac 331
          || ||  |||||||  ||||||| ||||||||||||||||||||| || || ||||||
Qy    258 CCCTCTGCTTCTGAGGAGCAGGTAGCCCAGGACACAGAGGAGGTTTTCCGCAGCTAC 314


RESULT   15
LOCUS       AA220617      331 bp    mRNA            EST       10-FEB-1997
```

```
DEFINITION  my25f11.r1 Barstead mouse pooled organs MPLRB4 Mus musculus cDNA
            clone 696909 5'.
ACCESSION   AA220617
NID         g1838404
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Mus.
REFERENCE   1  (bases 1 to 331)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT
            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:430469
            Seq primer: -28m13 rev2 ET from Amersham.
FEATURES             Location/Qualifiers
     source          1..331
                     /organism="Mus musculus"
                     /strain="FVB/N"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: EcoRI; Site_2: NotI; 1st strand cDNA
                     was primed with a Not I - oligo(dT) primer [5'
                     TGTTACGAATCTGAAGTGGGAGCGGCCGCCCTTTTTTTTTTTTTTTTTTTTTTT
                     3']; double-stranded cDNA was ligated to Eco RI adaptors
                     [GTTGGATTCGGTACC], digested with Not I and cloned into
                     the Not I and Eco RI sites of the modified pT7T3 vector.
                     Library constructed by Bob Barstead."
                     /clone="696909"
                     /clone_lib="Barstead mouse pooled organs MPLRB4"
                     /sex="mixed"
                     /dev_stage="7 day"
                     /lab_host="DH10B"
     mRNA            <1..>331
BASE COUNT       78 a    104 c     89 g     60 t
ORIGIN

  Query Match            1.9%;  Score 39;  DB 49;  Length 331;
  Best Local Similarity  84.2%;  Pred. No. 6.98e-27;
  Matches    48;  Conservative    0;  Mismatches   9;  Indels   0;  Gaps   0;

Db    275 ccgtccccttctgaacagcaggttgcccaggacacagaggaggtctttcgaagctac 331
          || ||  |||||||  ||||||| ||||||||||||||||||| || || ||||||
Qy    258 CCCTCTGCTTCTGAGGAGCAGGTAGCCCAGGACACAGAGGAGGTTTTCCGCAGCTAC 314


Search completed: Thu Aug 21 11:52:14 1997
Job time : 1336 secs.
```

****

MPSRCH (TM)

****

Release 2.1D John F. Collins, Biocomputing Research Unit.

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Wed Aug 20 10:59:48 1997; MasPar time 5.42 Seconds
445.716 Million cell updates/sec

Tabular output not generated.

Title: >US-08-320-157-7
Description: (1-211) from US08320157.pep
Perfect Score: 1561
Sequence: 1 MASGQGPGPPRQECGEPALP..........LVVLGVVLLGQFVVRRFFKS 211

Scoring table: PAM 150
Gap 11

Searched: 96640 seqs, 11439865 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: a-geneseq27
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20

Statistics: Mean 32.533; Variance 142.854; scale 0.228

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description	Pred. No.
1	1561	100.0	211	20	W03668	Bak protein.	5.68e-137
2	1561	100.0	211	17	R81451	Bcl-Y apoptosis-relat	5.68e-137
3	1561	100.0	211	13	R77876	Human Cdn-1.	5.68e-137
4	1529	98.0	211	20	W03669	Bak-2 protein.	7.28e-134
5	1529	98.0	211	13	R77877	Human Cdn-2.	7.28e-134
6	1154	73.9	152	13	R77879	Human Cdn-1(60-211).	1.48e-97
7	1079	69.1	141	13	R77880	Human Cdn-1(71-211).	2.53e-90
8	898	57.5	116	13	R77881	Human Cdn-1(96-211).	6.25e-73
9	274	17.6	190	13	R68884	Chicken lymphoid BCL-	1.74e-14
10	272	17.4	232	19	W01020	Apoptosis-blocking pr	2.62e-14
11	266	17.0	233	13	R68887	Human thymus BCL-XL.	8.94e-14
12	266	17.0	233	20	W05821	Bcl-XL protein.	8.94e-14
13	261	16.7	232	19	W01019	Apoptosis-blocking pr	2.48e-13
14	261	16.7	239	19	W01018	Apoptosis-blocking pr	2.48e-13
15	259	16.6	205	13	R71405	Human bcl-2 beta prot	3.72e-13
16	259	16.6	205	13	R68886	Human thymus BCL-2.	3.72e-13
17	259	16.6	205	13	R70332	Human bcl-2 protein.	3.72e-13
18	259	16.6	239	13	R71404	Human bcl-2 alpha pro	3.72e-13
19	259	16.6	239	9	R47344	Human oncogene bcl-2	3.72e-13
20	259	16.6	239	1	P80987	Sequence of bcl-2-alp	3.72e-13
21	259	16.6	239	13	R70331	Human bcl-2 protein.	3.72e-13
22	259	16.6	239	8	R42312	Bcl-2 oncogene produc	3.72e-13
23	229	14.7	205	1	P80988	Sequence of bcl-2-bet	1.59e-10
24	219	14.0	229	19	W01021	Apoptosis-blocking pr	1.17e-09
25	216	13.8	192	13	R71407	Murine Bax protein.	2.13e-09
26	213	13.6	88	13	R77878	Human Cdn-3.	3.87e-09
27	206	13.2	192	13	R71406	Human Bax protein.	1.54e-08
28	195	12.5	350	12	R68814	Human mcl-1 gene prod	1.34e-07
29	166	10.6	63	13	R68885	Human thymus BCL-X1.	3.63e-05
30	98	6.3	277	14	R76996	Deduced sequence enco	7.55e+00
31	97	6.2	170	13	R68888	Human thymus BCL-XS.	8.91e+00
32	94	6.0	338	14	R73955	Duffy blood group gpD	1.46e+01
33	89	5.7	277	15	R79904	ACT-4-h-1 receptor se	3.28e+01
34	89	5.7	277	14	R74737	ACT-4 cell surface re	3.28e+01
35	89	5.7	471	14	R79163	Partial sequence of b	3.28e+01
36	89	5.7	1157	9	R48678	Insecticidal protoxin	3.28e+01
37	89	5.7	1984	18	R99639	Peripheral nervous sy	3.28e+01
38	89	5.7	1989	18	R92317	Peripheral nervous sy	3.28e+01
39	87	5.6	451	14	R79909	N. crassa mtr gene pr	4.51e+01
40	87	5.6	470	9	R47118	Mtr protein of Neuros	4.51e+01
41	88	5.6	2749	3	R13887	Inositol-3-phosphate	3.85e+01
42	86	5.5	972	19	W06331	Full-length receptor-	5.28e+01
43	86	5.5	972	15	R76468	Mature non-differenti	5.28e+01
44	86	5.5	987	19	W06335	Full length receptor-	5.28e+01
45	86	5.5	987	15	R89263	Human non-differentia	5.28e+01

ALIGNMENTS

RESULT 1
ID W03668 standard; Protein; 211 AA.
AC W03668;
DT 22-FEB-1997 (first entry)
DE Bak protein.
KW Human; Bak; apoptosis; latency; virus replication;
KW Epstein-Barr virus; BHRF1; fusion protein; epitope tag;
KW drug screening; co-precipitation; ELISA; immunoassay; antibody;
KW protein interactive trapping; virucide; antitumour; diagnostic.
OS Homo sapiens.
PN WO9633416-A1.
PD 24-OCT-1996.
PF 19-APR-1996; U05639.
PR 20-APR-1995; US-426529.
PA (LXRB-) LXR BIOTECHNOLOGY INC.
PI Barr PJ, Kiefer MC;
DR WPI; 96-485886/48.
DR N-PSDB; T42138.
PT Screening for anti-viral agents - by detecting the ability of an
PT agent to disrupt the interaction of a Bak protein and a viral
PT protein
PS Disclosure; Fig 1; 24pp; English.
CC This Bak protein sequence represents a bcl-1 homologue which
CC interacts with Epstein-Barr virus (EBV) early lytic cycle BHRF1
CC protein, and is capable of modulating apoptosis. The protein may
CC be used in complete or partial form, or as an epitope tag fusion
CC protein, in a new virucide drug screening method, which involves
CC combination of Bak protein and a viral protein (e.g. EBV BHRF1),
CC exposure to a test compound, and monitoring for disruption of the
CC interaction, e.g. by co-precipitation, protein interactive trapping
CC or ELISA. Interaction of Bak and viral proteins allows viral
CC replication or latency in the absence of apoptosis. Compounds which
CC inhibit the interaction may be used as virucide, antitumour or
CC diagnostic agents.
SQ Sequence 211 AA;

B.D.

```
  Query Match          100.0%;  Score 1561;  DB 20;  Length 211;
  Best Local Similarity 100.0%;  Pred. No. 5.68e-137;
  Matches  211;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Db       1 masgqgpgpprqecgepalpsaseeqvaqdteevfrsyvfyrhqqeqeaegvaapadpem 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       1 MASGQGPGPPRQECGEPALPSASEEQVAQDTEEVFRSYVFYRHQQEQEAEGVAAPADPEM 60
```

```
Db     61 vtlplqpsstmgqvgrqlaiigddinrrydsefqtmlqhlqptaenayeyftkiatslfe 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     61 VTLPLQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENAYEYFTKIATSLFE 120

Db    121 sginwgrvvallgfgyrlalhvyqhgltgflgqvtrfvvdfmlhhciarwiaqrggwvaa 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    121 SGINWGRVVALLGFGYRLALHVYQHGLTGFLGQVTRFVVDFMLHHCIARWIAQRGGWVAA 180

Db    181 lnlgngpilnvlvvlgvvllgqfvvrrffks 211
          |||||||||||||||||||||||||||||||
Qy    181 LNLGNGPILNVLVVLGVVLLGQFVVRRFFKS 211


RESULT    2
ID    R81451 standard; Protein; 211 AA.
AC    R81451;
DT    02-JUL-1996  (first entry)
DE    Bcl-Y apoptosis-related protein.
KW    Bcl-Y; apoptosis; cell proliferation; cell death; diagnosis;
KW    therapy.
OS    Homo sapiens.
FH    Key             Location/Qualifiers
FT    Domain          188..205
FT    /label= C-terminal_domain
FT    /note= "putative membrane localisation sequence"
PN    WO9605232-A1.
PD    22-FEB-1996.
PF    09-AUG-1995; U10103.
PR    09-AUG-1994; US-287427.
PR    11-OCT-1994; US-321071.
PA    (IMMU-) IMMUNOGEN INC.
PI    Chittenden TD;
DR    WPI; 96-139648/14.
DR    N-PSDB; T17375.
PT    New isolated human Bcl-Y protein - used to develop prods. for
PT    treating disorders characterised by inappropriate cell proliferation
PT    or cell death
PS    Claim 3; Fig 4; 100pp; English.
CC    Bcl-Y protein (R81451) is a member of the Bcl-2 family and can
CC    induce apoptosis in cells and function as a negative regulator of
CC    Bcl-2 function.  Bcl-Y mRNA was detected in all human tumour cell
CC    lines examined and is also widely expressed in primary human
CC    tissues.  It can be obtd. by expression of a full-length cDNA
CC    clone (T17375) in pref. mammalian host cells.  Bcl-Y can be used to
CC    develop prods. for treating disorders associated with inappropriate
CC    cell proliferation or cell death, and to raise antibodies used for
CC    the diagnosis or monitoring of such disorders.
SQ    Sequence   211 AA;

  Query Match          100.0%;  Score 1561;  DB 17;  Length 211;
  Best Local Similarity 100.0%;  Pred. No. 5.68e-137;
  Matches   211;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      1 masgqgpgpprqecgepalpsaseeqvaqdteevfrsyvfyrhqqeqeaegvaapadpem 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      1 MASGQGPGPPRQECGEPALPSASEEQVAQDTEEVFRSYVFYRHQQEQEAEGVAAPADPEM 60

Db     61 vtlplqpsstmgqvgrqlaiigddinrrydsefqtmlqhlqptaenayeyftkiatslfe 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     61 VTLPLQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENAYEYFTKIATSLFE 120

Db    121 sginwgrvvallgfgyrlalhvyqhgltgflgqvtrfvvdfmlhhciarwiaqrggwvaa 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    121 SGINWGRVVALLGFGYRLALHVYQHGLTGFLGQVTRFVVDFMLHHCIARWIAQRGGWVAA 180

Db    181 lnlgngpilnvlvvlgvvllgqfvvrrffks 211
          |||||||||||||||||||||||||||||||
Qy    181 LNLGNGPILNVLVVLGVVLLGQFVVRRFFKS 211


RESULT    3
ID    R77876 standard; Protein; 211 AA.
AC    R77876;
DT    21-NOV-1995  (first entry)
DE    Human Cdn-1.
KW    Cdn-1; apoptosis modulator; adoptive immunotherapy; therapy; HIV;
KW    autoimmune disease; reperfusion injury; hepatitis, osteoporosis;
KW    shock; lymphoma; eczema.
OS    Homo sapiens.
PN    WO9515084-A.
PD    08-JUN-1995.
PF    30-NOV-1994; U13930.
PR    30-NOV-1993; US-160067.
PR    07-OCT-1994; US-320157.
PA    (LXRB-) LXR BIOTECHNOLOGY INC.
PI    Barr PJ,  Kiefer MC;
DR    WPI; 95-215106/28.
DR    N-PSDB; Q95492.
PT    New nucleic acid sequences encoding Cdn apoptosis modulators  - and
PT    related vectors, transformed cells, proteins and antibodies, useful
PT    or diagnosis and treatment e.g. of HIV infection, reperfusion injury
PT    etc.
PS    Disclosure; Fig.3A-B; 66pp; English.
CC    Cdn-1 cDNA was isolated from a human heart cDNA library using a
CC    previously isolated clone as probe.  Recombinant Cdn-1 was produced
CC    in Sf9 and human colon adenocarcinoma HT29 cells.  Expression of
CC    Cdn-1 in WI-L2 lymphoblastoid cells resulted in increased cell
CC    survival in response to anti-Fas-mediated apoptosis.
SQ    Sequence   211 AA;

  Query Match          100.0%;  Score 1561;  DB 13;  Length 211;
  Best Local Similarity 100.0%;  Pred. No. 5.68e-137;
  Matches   211;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      1 masgqgpgpprqecgepalpsaseeqvaqdteevfrsyvfyrhqqeqeaegvaapadpem 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      1 MASGQGPGPPRQECGEPALPSASEEQVAQDTEEVFRSYVFYRHQQEQEAEGVAAPADPEM 60

Db     61 vtlplqpsstmgqvgrqlaiigddinrrydsefqtmlqhlqptaenayeyftkiatslfe 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     61 VTLPLQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENAYEYFTKIATSLFE 120

Db    121 sginwgrvvallgfgyrlalhvyqhgltgflgqvtrfvvdfmlhhciarwiaqrggwvaa 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    121 SGINWGRVVALLGFGYRLALHVYQHGLTGFLGQVTRFVVDFMLHHCIARWIAQRGGWVAA 180

Db    181 lnlgngpilnvlvvlgvvllgqfvvrrffks 211
          |||||||||||||||||||||||||||||||
Qy    181 LNLGNGPILNVLVVLGVVLLGQFVVRRFFKS 211


RESULT    4
ID    W03669 standard; Protein; 211 AA.
AC    W03669;
DT    22-FEB-1997  (first entry)
DE    Bak-2 protein.
KW    Human; Bak-2; apoptosis; latency; virus replication;
KW    Epstein-Barr virus; BHRF1; fusion protein; epitope tag;
KW    drug screening; co-precipitation; ELISA; immunoassay; antibody;
KW    protein interactive trapping; virucide; antitumour; diagnostic.
OS    Homo sapiens.
PN    WO9633416-A1.
PD    24-OCT-1996.
PF    19-APR-1996; U05639.
PR    20-APR-1995; US-426529.
PA    (LXRB-) LXR BIOTECHNOLOGY INC.
PI    Barr PJ,  Kiefer MC;
DR    WPI; 96-485886/48.
DR    N-PSDB; T42139.
PT    Screening for anti-viral agents - by detecting the ability of an
PT    agent to disrupt the interaction of a Bak protein and a viral
PT    protein
PS    Disclosure; Fig 2; 24pp; English.
CC    This Bak-2 protein sequence represents a bcl-1 homologue which
CC    interacts with Epstein-Barr virus (EBV) early lytic cycle BHRF1
CC    protein, and is capable of modulating apoptosis.  The protein may
```

```
CC   be used in complete or partial form, or as an epitope tag fusion
CC   protein, in a new virucide drug screening method, which involves
CC   combination of Bak-2 protein and a viral protein (e.g. EBV BHRF1),
CC   exposure to a test compound, and monitoring for disruption of the
CC   interaction, e.g. by co-precipitation, protein interactive trapping
CC   or ELISA.  Interaction of Bak-2 and viral proteins allows viral
CC   replication or latency in the absence of apoptosis.  Compounds which
CC   inhibit the interaction may be used as virucide, antitumour or
CC   diagnostic agents.
SQ   Sequence   211 AA;

  Query Match           98.0%;  Score 1529;  DB 20;  Length 211;
  Best Local Similarity  97.2%;  Pred. No. 7.28e-134;
  Matches   205;  Conservative    5;  Mismatches   1;  Indels   0;  Gaps   0;

Db      1 masgqgpgpprqecgepalpsaseeqvaqdteevfrsyvfyhhqqeqeaegaaapadpem 60
          ||||||||||||||||||||||||||||||||||||||||:|||||||| ||||||||
Qy      1 MASGQGPGPPRQECGEPALPSASEEQVAQDTEEVFRSYVFYRHQQEQEAEGVAAPADPEM 60

Db     61 vtlplqpsstmgqvgrqlaiigddinrrydsefqtmlqhlqptaenayeyftkiasslfe 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||
Qy     61 VTLPLQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENAYEYFTKIATSLFE 120

Db    121 sginwgrvvallgfsyrlalhiyqrgltgflgqvtrfvvdfmlhhciarwiaqrggwvaa 180
          ||||||||||||||:||||||:||:||||||||||||||||||||||||||||||||||
Qy    121 SGINWGRVVALLGFGYRLALHVYQHGLTGFLGQVTRFVVDFMLHHCIARWIAQRGGWVAA 180

Db    181 lnlgngpilnvlvvlgvvllgqfvvrrffks 211
          |||||||||||||||||||||||||||||||
Qy    181 LNLGNGPILNVLVVLGVVLLGQFVVRRFFKS 211


RESULT    5
ID   R77877 standard; Protein; 211 AA.
AC   R77877;
DT   21-NOV-1995  (first entry)
DE   Human Cdn-2.
KW   Cdn-2; apoptosis modulator; adoptive immunotherapy; therapy; HIV;
KW   autoimmune disease; reperfusion injury; hepatitis, osteoporosis;
KW   shock; lymphoma; eczema.
OS   Homo sapiens.
PN   WO9515084-A.
PD   08-JUN-1995.
PF   30-NOV-1994; U13930.
PR   30-NOV-1993; US-160067.
PR   07-OCT-1994; US-320157.
PA   (LXRB-) LXR BIOTECHNOLOGY INC.
PI   Barr PJ,  Kiefer MC;
DR   WPI; 95-215106/28.
DR   N-PSDB; Q95493.
PT   New nucleic acid sequences encoding Cdn apoptosis modulators  - and
PT   related vectors, transformed cells, proteins and antibodies, useful
PT   or diagnosis and treatment e.g. of HIV infection, reperfusion injury
PT   etc.
PS   Disclosure; Fig.5D-E; 66pp; English.
CC   Cdn-2 cDNA was isold. from a human placental genomic library
CC   using a 950 bp fragment of Cdn-1 cDNA.  Expression of Cdn-2
CC   in mouse progenitor B-cell FL5.12 cells decreased IL-3-induced
CC   apoptosis.  The Cdn-2 protein displayed 97% amino acid identity
CC   with Cdn-1 (R77876).
SQ   Sequence   211 AA;

  Query Match           98.0%;  Score 1529;  DB 13;  Length 211;
  Best Local Similarity  97.2%;  Pred. No. 7.28e-134;
  Matches   205;  Conservative    5;  Mismatches   1;  Indels   0;  Gaps   0;

Db      1 masgqgpgpprqecgepalpsaseeqvaqdteevfrsyvfyhhqqeqeaegaaapadpem 60
          ||||||||||||||||||||||||||||||||||||||||:|||||||| ||||||||
Qy      1 MASGQGPGPPRQECGEPALPSASEEQVAQDTEEVFRSYVFYRHQQEQEAEGVAAPADPEM 60

Db     61 vtlplqpsstmgqvgrqlaiigddinrrydsefqtmlqhlqptaenayeyftkiasslfe 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||
Qy     61 VTLPLQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENAYEYFTKIATSLFE 120

Db    121 sginwgrvvallgfsyrlalhiyqrgltgflgqvtrfvvdfmlhhciarwiaqrggwvaa 180
          ||||||||||||||:||||||:||:||||||||||||||||||||||||||||||||||
Qy    121 SGINWGRVVALLGFGYRLALHVYQHGLTGFLGQVTRFVVDFMLHHCIARWIAQRGGWVAA 180

Db    181 lnlgngpilnvlvvlgvvllgqfvvrrffks 211
          |||||||||||||||||||||||||||||||
Qy    181 LNLGNGPILNVLVVLGVVLLGQFVVRRFFKS 211


RESULT    6
ID   R77879 standard; Protein; 152 AA.
AC   R77879;
DT   21-NOV-1995  (first entry)
DE   Human Cdn-1(60-211).
KW   Cdn-1; apoptosis modulator; adoptive immunotherapy; therapy; HIV;
KW   autoimmune disease; reperfusion injury; hepatitis, osteoporosis;
KW   shock; lymphoma; eczema.
OS   Homo sapiens.
PN   WO9515084-A.
PD   08-JUN-1995.
PF   30-NOV-1994; U13930.
PR   30-NOV-1993; US-160067.
PR   07-OCT-1994; US-320157.
PA   (LXRB-) LXR BIOTECHNOLOGY INC.
PI   Barr PJ,  Kiefer MC;
DR   WPI; 95-215106/28.
PT   New nucleic acid sequences encoding Cdn apoptosis modulators  - and
PT   related vectors, transformed cells, proteins and antibodies, useful
PT   or diagnosis and treatment e.g. of HIV infection, reperfusion injury
PT   etc.
PS   Disclosure; Fig.11; 66pp; English.
CC   Expression of Cdn-1 in WI-L2 lymphoblastoid cells resulted in
CC   increased cell survival in response to anti-Fas-mediated apoptosis.
CC   Deletion of the N-terminal 59 amino acids of Cdn-1 only slightly
CC   decreased this activity, suggesting that small, truncated Cdn-1
CC   molecules may be potent therapeutics.
SQ   Sequence   152 AA;

  Query Match           73.9%;  Score 1154;  DB 13;  Length 152;
  Best Local Similarity 100.0%;  Pred. No. 1.48e-97;
  Matches   152;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      1 mvtlplqpsstmgqvgrqlaiigddinrrydsefqtmlqhlqptaenayeyftkiatslf 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     60 MVTLPLQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENAYEYFTKIATSLF 119

Db     61 esginwgrvvallgfgyrlalhvyqhgltgflgqvtrfvvdfmlhhciarwiaqrggwva 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    120 ESGINWGRVVALLGFGYRLALHVYQHGLTGFLGQVTRFVVDFMLHHCIARWIAQRGGWVA 179

Db    121 alnlgngpilnvlvvlgvvllgqfvvrrffks 152
          |||||||||||||||||||||||||||||||
Qy    180 ALNLGNGPILNVLVVLGVVLLGQFVVRRFFKS 211


RESULT    7
ID   R77880 standard; Protein; 141 AA.
AC   R77880;
DT   21-NOV-1995  (first entry)
DE   Human Cdn-1(71-211).
KW   Cdn-1; apoptosis modulator; adoptive immunotherapy; therapy; HIV;
KW   autoimmune disease; reperfusion injury; hepatitis, osteoporosis;
KW   shock; lymphoma; eczema.
OS   Homo sapiens.
PN   WO9515084-A.
PD   08-JUN-1995.
PF   30-NOV-1994; U13930.
PR   30-NOV-1993; US-160067.
PR   07-OCT-1994; US-320157.
PA   (LXRB-) LXR BIOTECHNOLOGY INC.
PI   Barr PJ,  Kiefer MC;
DR   WPI; 95-215106/28.
```

```
PT   New nucleic acid sequences encoding Cdn apoptosis modulators - and
PT   related vectors, transformed cells, proteins and antibodies, useful
PT   or diagnosis and treatment e.g. of HIV infection, reperfusion injury
PT   etc.
PS   Disclosure; Fig.11; 66pp; English.
CC   Expression of Cdn-1 in WI-L2 lymphoblastoid cells resulted in
CC   increased cell survival in response to anti-Fas-mediated apoptosis.
CC   Deletion of the N-terminal 70 amino acids of Cdn-1 improved this
CC   activity, suggesting that small, truncated Cdn-1 molecules may be
CC   potent therapeutics.
SQ   Sequence  141 AA;

  Query Match           69.1%;  Score 1079;  DB 13;  Length 141;
  Best Local Similarity 100.0%;  Pred. No. 2.53e-90;
  Matches  141;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      1 mgqvgrqlaiigddinrrydsefqtmlqhlqptaenayeyftkiatslfesginwgrvva 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     71 MGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENAYEYFTKIATSLFESGINWGRVVA 130

Db     61 llgfgyrlalhvyqhgltgflgqvtrfvvdfmlhhciarwiaqrggwvaalnlgngpiln 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    131 LLGFGYRLALHVYQHGLTGFLGQVTRFVVDFMLHHCIARWIAQRGGWVAALNLGNGPILN 190

Db    121 vlvvlgvvllgqfvvrrffks 141
          |||||||||||||||||||||
Qy    191 VLVVLGVVLLGQFVVRRFFKS 211


RESULT   8
ID   R77881 standard; Protein; 116 AA.
AC   R77881;
DT   21-NOV-1995  (first entry)
DE   Human Cdn-1(96-211).
KW   Cdn-1; apoptosis modulator; adoptive immunotherapy; therapy; HIV;
KW   autoimmune disease; reperfusion injury; hepatitis, osteoporosis;
KW   shock; lymphoma; eczema.
OS   Homo sapiens.
PN   WO9515084-A.
PD   08-JUN-1995.
PF   30-NOV-1994; U13930.
PR   30-NOV-1993; US-160067.
PR   07-OCT-1994; US-320157.
PA   (LXRB-) LXR BIOTECHNOLOGY INC.
PI   Barr PJ,  Kiefer MC;
DR   WPI; 95-215106/28.
PT   New nucleic acid sequences encoding Cdn apoptosis modulators - and
PT   related vectors, transformed cells, proteins and antibodies, useful
PT   or diagnosis and treatment e.g. of HIV infection, reperfusion injury
PT   etc.
PS   Disclosure; Fig.11; 66pp; English.
CC   Expression of Cdn-1 in WI-L2 lymphoblastoid cells resulted in
CC   increased cell survival in response to anti-Fas-mediated apoptosis.
CC   Truncated Cdn-1 derivatives given in R77879-81 were used to
CC   test the effects of deleting the N-terminal sequences of Cdn-1
CC   on this activity.
SQ   Sequence  116 AA;

  Query Match           57.5%;  Score 898;  DB 13;  Length 116;
  Best Local Similarity 100.0%;  Pred. No. 6.25e-73;
  Matches  116;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      1 mlqhlqptaenayeyftkiatslfesginwgrvvallgfgyrlalhvyqhgltgflgqvt 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     96 MLQHLQPTAENAYEYFTKIATSLFESGINWGRVVALLGFGYRLALHVYQHGLTGFLGQVT 155

Db     61 rfvvdfmlhhciarwiaqrggwvaalnlgngpilnvlvvlgvvllgqfvvrrffks 116
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    156 RFVVDFMLHHCIARWIAQRGGWVAALNLGNGPILNVLVVLGVVLLGQFVVRRFFKS 211


RESULT   9
ID   R68884 standard; Protein; 190 AA.
AC   R68884;
DT   10-AUG-1995  (first entry)
DE   Chicken lymphoid BCL-X.
KW   Chicken; bird; fowl; BCL-X; apoptosis; cell death; cancer;
KW   neurodegenerative disease; autoimmune disease; Parkinson's disease;
KW   amylotrophic lateral sclerosis; multiple sclerosis; oncogene.
OS   Gallus domesticus.
PN   WO9500642-A.
PD   05-JAN-1995.
PF   22-JUN-1994; U07089.
PR   22-JUN-1993; US-081448.
PA   (ARCH-) ARCH DEV CORP.
PA   (UNMI ) UNIV MICHIGAN.
PI   Boise LH,  Nunez G,  Thompson CB;
DR   WPI; 95-052079/07.
DR   N-PSDB; Q81696.
PT   New poly-nucleotide encoding new poly-peptide(s) that modify
PT   apoptosis - and related vectors, recombinant cells and
PT   antibodies, useful in assay and for control of cell death in e.g.
PT   neuronal cells, lymphocytes and cancers
PS   Claim 4; Page 87; 127pp; English.
CC   This protein may be expressed recombinantly, particularly with pcmV
CC   plasmids as vectors for expression in mammalian cell cultures.
CC   The protein has particular application in cancer cells (failure of
CC   programmed cell death (PCD)) or neurodegenerative and autoimmune diseases
CC   (premature PCD), e.g. Parkinson's disease, amylotrophic lateral
CC   sclerosis and multiple sclerosis.
SQ   Sequence  190 AA;

  Query Match           17.6%;  Score 274;  DB 13;  Length 190;
  Best Local Similarity 28.2%;  Pred. No. 1.74e-14;
  Matches   37;  Conservative   33;  Mismatches  59;  Indels   2;  Gaps   2;

Db     60 vvngatvhrsslevheivrasdvrqalrdagdefelryrrafsdltsqlhitpgtayqsf 119
          |   |     :| ::     ::| : |   ||::: ||   |  :  :|: |:  ||: |
Qy     52 VAAPADPEMVTLPLQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENAYEYF 111

Db    120 eqvvnelfhdgvnwgrivaffsfggalcvesvdkemrvlvgrivswmttyltdh-ldpwi 178
            :   ||  |:||||:||:::||  | :   :  :  ::|::   :  ::  | :  ||
Qy    112 TKIATSLFESGINWGRVVALLGFGYRLALHVYQHGLTGFLGQVTRFVVDFMLHHCIARWI 171

Db    179 qenggwvrtal 189
            : |||| :||
Qy    172 AQRGGWV-AAL 181


RESULT   10
ID   W01020 standard; Protein; 232 AA.
AC   W01020;
DT   18-DEC-1996  (first entry)
DE   Apoptosis-blocking protein Bcl-2 mutant 80-6 (del80-86).
KW   Apoptosis-regulating protein; Bcl-2; oncogene;
KW   adenovirus E1B 19K protein; cell death; cancer; tumour;
KW   immune disorder; diagnosis; therapy; Bip1A; Bip13; Bip5; NIp1;
KW   Nip2; Nip3.
OS   Synthetic.
PN   EP-733706-A2.
PD   25-SEP-1996.
PF   21-MAR-1996; 104542.
PR   21-MAR-1995; US-408095.
PA   (UYSL-) UNIV ST LOUIS.
PI   Chinnadurai G;
DR   WPI; 96-427055/43.
PT   Nucleic acids encoding apoptosis regulating proteins - useful for
PT   diagnosing and treating immune disorders, malignancies, etc.
PS   Example 8; Page 34-35; 60pp; English.
CC   The 80-6 mutant (W01020) of the bcl-2 oncogene product (W01018)
CC   lacks amino acids 80-86 of the native protein.  This and other
CC   Bcl-2 mutants (see also W01019-21) were used in a two hybrid assay
CC   to examine the interactions between Bcl-2 and novel apoptosis-
CC   regulating proteins Nip1, Nip2 and Nip3 (W00997-99).  2 Motifs
CC   (W01003-04) on Bcl-2 were identified that are essential for
CC   interaction with the Nip proteins.  These motifs show homology
```

```
CC   to motifs (W01005-06) identified on the adenovirus E1B 19K
CC   apoptosis-blocking protein (W01010).
SQ   Sequence   232 AA;

  Query Match           17.4%;  Score 272;  DB 19;  Length 232;
  Best Local Similarity  29.7%;  Pred. No. 2.62e-14;
  Matches    47;  Conservative   41;  Mismatches  62;  Indels   8;  Gaps   6;

Db     48 ffssqpghtphp-aasrdpvartsplqtpaapgspvppvvhltlrqagddfsrryrrdfa 106
          |:  |  : :   ||: ||   | ||| ::: |  |    :|::   |||::|||  :|
Qy     40 FYRHQQEQEAEGVAAPADPEMVTLPLQPSSTMG-QVGR--QLAI--IGDDINRRYDSEFQ 94

Db    107 emssqlhltpftargcfatvveelfrdgvnwgrivaffefggvmcvesvnremsplvdni 166
           |  :|: |:  |   |: :   ||  |:||||:||:: ||  : :    : :: ::  :
Qy     95 TMLQHLQPTAENAYEYFTKIATSLFESGINWGRVVALLGFGYRLALHVYQHGLTGFLGQV 154

Db    167 alwmteylnrh-lhtwiqdnggwdafvel-ygpsmrpl 202
          :  : ::: :| :  || : ||| | ::|  || :  |
Qy    155 TRFVVDFMLHHCIARWIAQRGGWVAALNLGNGPILNVL 192


RESULT   11
ID   R68887 standard; Protein; 233 AA.
AC   R68887;
DT   10-AUG-1995  (first entry)
DE   Human thymus BCL-XL.
KW   BCL-XL; apoptosis; cell death; cancer; neurodegenerative disease;
KW   autoimmune disease; Parkinson disease; amylotrophic lateral sclerosis;
KW   multiple sclerosis.
OS   Homo sapiens.
PN   WO9500642-A.
PD   05-JAN-1995.
PF   22-JUN-1994; U07089.
PR   22-JUN-1993; US-081448.
PA   (ARCH-) ARCH DEV CORP.
PA   (UNMI ) UNIV MICHIGAN.
PI   Boise LH,  Nunez G,  Thompson CB;
DR   WPI; 95-052079/07.
DR   N-PSDB; Q81698.
PT   New poly-nucleotide encoding new poly-peptide(s) that modify
PT   apoptosis - and related vectors, recombinant cells and
PT   antibodies, useful in assay and for control of cell death in e.g.
PT   neuronal cells, lymphocytes and cancers
PS   Claim 3; Page 94; 127pp; English.
CC   This protein may be expressed recombinantly, particularly with pcmV
CC   plasmids as vectors for expression in mammalian cell cultures.  The
CC   protein has particular application in cancer cells (failure of
CC   programmed cell death (PCD)) or neurodegenerative and autoimmune
CC   diseases (premature PCD), e.g. Parkinson's disease, amylotrophic
CC   lateral sclerosis and multiple sclerosis.
SQ   Sequence   233 AA;

  Query Match           17.0%;  Score 266;  DB 13;  Length 233;
  Best Local Similarity  25.5%;  Pred. No. 8.94e-14;
  Matches    36;  Conservative   42;  Mismatches  60;  Indels   3;  Gaps   3;

Db     61 dspavngatghs-ssldarevipmaavkqalreagdefelryrrafsdltsqlhitpgta 119
          :: :| :::     :|  :    |: | : |   ||::: ||   |  :  :|: |:  |
Qy     48 EAEGVAAPADPEMVTLPLQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENA 107

Db    120 yqsfeqvvnelfrdgvnwgrivaffsfggalcvesvdkemqvlvsriaawmatylndh-l 178
          |: |  :   ||  |:||||:||:::||  | :   :  :  :::::  :  ::  | :
Qy    108 YEYFTKIATSLFESGINWGRVVALLGFGYRLALHVYQHGLTGFLGQVTRFVVDFMLHHCI 167

Db    179 epwiqenggwdtfvelygnna 199
            || : ||| : ::| || :
Qy    168 ARWIAQRGGWVAALNL-GNGP 187


RESULT   12
ID   W05821 standard; Protein; 233 AA.
AC   W05821;
DT   30-MAR-1997  (first entry)
DE   Bcl-XL protein.
KW   Human; bcl-XL; T-lymphocyte; cell death; BH1 domain; BH2 domain;
KW   Bcl-2 homology domain; gene therapy; HIV; AIDS; antisense;
KW   immune disorder; autoimmune disease; graft rejection;
KW   graft-versus-host disease; apoptosis; adoptive immunotherapy.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   Domain          129..148
FT   /note= "BH1 domain"
FT   Domain          180..191
FT   /note= "BH2 domain"
PN   WO9634956-A1.
PD   07-NOV-1996.
PF   02-MAY-1996; U06203.
PR   04-MAY-1995; US-435518.
PR   07-JUN-1995; US-481739.
PA   (ARCH-) ARCH DEV CORP.
PA   (USNA ) US SEC OF NAVY.
PI   June CH,  Thompson CB;
DR   WPI; 96-506159/50.
DR   N-PSDB; T40079.
PT   Inducing or preventing death of T cells by bcl-XL protein regulation
PT   - used to increase survival of HIV infected cells or to
PT   down:regulate immune responses in immune diseases
PS   Disclosure; Page 52-53; 76pp; English.
CC   This is the sequence of a human bcl-XL protein, which protects
CC   T-lymphocytes against cell death.  A splice variant form, bcl-XS,
CC   lacks a stretch of 63 amino acids, and is a dominant negative
CC   regulator of bcl-XL function.  The gene may be modified to
CC   facilitate interaction with costimulatory Bax protein and inhibit
CC   interaction with antagonistic Bad protein, by modification of the
CC   Bcl-2 homology domains BH1 and/or BH2.  The bcl-XL gene may be
CC   introduced into T-cells in vivo or ex vivo via gene transfer using
CC   a vector for HIV infection gene therapy, to augment intracellular
CC   bcl-XL protein levels and protect from cell death.  A corresponding
CC   antisense oligonucleotide or expression vector may be used in gene
CC   therapy of e.g. autoimmune disease, graft rejection or graft-
CC   versus-host disease, to induce cell death (e.g. apoptosis) and
CC   down-regulate the immune response in a T-lymphocyte population.
SQ   Sequence   233 AA;

  Query Match           17.0%;  Score 266;  DB 20;  Length 233;
  Best Local Similarity  25.5%;  Pred. No. 8.94e-14;
  Matches    36;  Conservative   42;  Mismatches  60;  Indels   3;  Gaps   3;

Db     61 dspavngatghs-ssldarevipmaavkqalreagdefelryrrafsdltsqlhitpgta 119
          :: :| :::     :|  :    |: | : |   ||::: ||   |  :  :|: |:  |
Qy     48 EAEGVAAPADPEMVTLPLQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENA 107

Db    120 yqsfeqvvnelfrdgvnwgrivaffsfggalcvesvdkemqvlvsriaawmatylndh-l 178
          |: |  :   ||  |:||||:||:::||  | :   :  :  :::::  :  ::  | :
Qy    108 YEYFTKIATSLFESGINWGRVVALLGFGYRLALHVYQHGLTGFLGQVTRFVVDFMLHHCI 167

Db    179 epwiqenggwdtfvelygnna 199
            || : ||| : ::| || :
Qy    168 ARWIAQRGGWVAALNL-GNGP 187


RESULT   13
ID   W01019 standard; Protein; 232 AA.
AC   W01019;
DT   18-DEC-1996  (first entry)
DE   Apoptosis-blocking protein Bcl-2 mutant 42-8 (del42-48).
KW   Apoptosis-regulating protein; Bcl-2; oncogene;
KW   adenovirus E1B 19K protein; cell death; cancer; tumour;
KW   immune disorder; diagnosis; therapy; Bip1A; Bip13; Bip5; NIp1;
KW   Nip2; Nip3.
OS   Synthetic.
PN   EP-733706-A2.
PD   25-SEP-1996.
PF   21-MAR-1996; 104542.
PR   21-MAR-1995; US-408095.
PA   (UYSL-) UNIV ST LOUIS.
```

```
PI   Chinnadurai G;
DR   WPI; 96-427055/43.
PT   Nucleic acids encoding apoptosis regulating proteins - useful for
PT   diagnosing and treating immune disorders, malignancies, etc.
PS   Example 8; Page 33-34; 60pp; English.
CC   The 42-8 mutant (W01019) of the bcl-2 oncogene product (W01018)
CC   lacks amino acids 42-48 of the native protein.  This and other
CC   Bcl-2 mutants (see also W01020-21) were used in a two hybrid assay
CC   to examine the interactions between Bcl-2 and novel apoptosis-
CC   regulating proteins Nip1, Nip2 and Nip3 (W00997-99).  The Nip
CC   proteins were unable to interact with mutant 42-8.  The site of
CC   deletion in this mutant corresponds to a motif (see also W01003)
CC   on Bcl-2 essential for interaction with Nip proteins.  A second
CC   binding motif (W01004) of Bcl-2 was also identified, and both
CC   show homology to motifs (W01005-06) found on the 19K protein
CC   (W01010) of adenovirus E1B.
SQ   Sequence   232 AA;

  Query Match           16.7%;  Score 261;  DB 19;  Length 232;
  Best Local Similarity  29.3%;  Pred. No. 2.48e-13;
  Matches    41;  Conservative   37;  Mismatches  57;  Indels   5;  Gaps   4;

Db     67 tpaapgaaagpalspvppvvhlt--lrqagddfsrryrrdfaemssqlhltpftargcfa 124
          :|| |   : | | | : : ::   |   |||::|||  :|  |  :|: |:  |   |:
Qy     54 APADPEMVTLP-LQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENAYEYFT 112

Db    125 tvveelfrdgvnwgrivaffefggvmcvesvnremsplvdnialwmteylnrh-lhtwiq 183
          :   ||  |:||||:||:: ||  : :    : :: ::  ::  : ::: :| :  ||
Qy    113 KIATSLFESGINWGRVVALLGFGYRLALHVYQHGLTGFLGQVTRFVVDFMLHHCIARWIA 172

Db    184 dnggwdafvel-ygpsmrpl 202
          : ||| | ::|  || :  |
Qy    173 QRGGWVAALNLGNGPILNVL 192


RESULT   14
ID   W01018 standard; Protein; 239 AA.
AC   W01018;
DT   18-DEC-1996  (first entry)
DE   Apoptosis-blocking protein Bcl-2.
KW   Apoptosis-regulating protein; Bcl-2; oncogene;
KW   adenovirus E1B 19K protein; cell death; cancer; tumour;
KW   immune disorder; diagnosis; therapy; Bip1A; Bip13; Bip5; NIp1;
KW   Nip2; Nip3.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   Binding_site    43..51
FT   /label= Binding_motif
FT   /note= "interacts with Bip proteins"
FT   Binding_site    106..112
FT   /label= Binding_motif
FT   /note= "interacts with Bip proteins"
PN   EP-733706-A2.
PD   25-SEP-1996.
PF   21-MAR-1996; 104542.
PR   21-MAR-1995; US-408095.
PA   (UYSL-) UNIV ST LOUIS.
PI   Chinnadurai G;
DR   WPI; 96-427055/43.
PT   Nucleic acids encoding apoptosis regulating proteins - useful for
PT   diagnosing and treating immune disorders, malignancies, etc.
PS   Example 8; Page 32-33; 60pp; English.
CC   The bcl-2 oncogene product (W01018) enhances the survival of
CC   haematopoietic B and T cells by blocking apoptosis induced by
CC   diverse agents.  Its activity is similar to that of the 19K
CC   protein (W01010) of adenovirus E1B.  3 Novel proteins, Bip1A,
CC   Bip13 and Bip5 (W01000-02), that specifically interact with
CC   Bcl-2, have been identified.  Mutational analysis (see also
CC   W01019-21) shows the apoptosis-regulating Nip proteins (see also
CC   W00997-99) associate with Bcl-2 at specific sites (see also
CC   W01003-04) that show homology to motifs (W01005-06) on 19K.
SQ   Sequence   239 AA;

  Query Match           16.7%;  Score 261;  DB 19;  Length 239;
  Best Local Similarity  29.3%;  Pred. No. 2.48e-13;
  Matches    41;  Conservative   37;  Mismatches  57;  Indels   5;  Gaps   4;

Db     74 tpaapgaaagpalspvppvvhlt--lrqagddfsrryrrdfaemssqlhltpftargcfa 131
          :|| |   : | | | : : ::   |   |||::|||  :|  |  :|: |:  |   |:
Qy     54 APADPEMVTLP-LQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENAYEYFT 112

Db    132 tvveelfrdgvnwgrivaffefggvmcvesvnremsplvdnialwmteylnrh-lhtwiq 190
          :   ||  |:||||:||:: ||  : :    : :: ::  ::  : ::: :| :  ||
Qy    113 KIATSLFESGINWGRVVALLGFGYRLALHVYQHGLTGFLGQVTRFVVDFMLHHCIARWIA 172

Db    191 dnggwdafvel-ygpsmrpl 209
          : ||| | ::|  || :  |
Qy    173 QRGGWVAALNLGNGPILNVL 192


RESULT   15
ID   R71405 standard; protein; 205 AA.
AC   R71405;
DT   30-OCT-1995  (first entry)
DE   Human bcl-2 beta protein.
KW   Human; bcl-2; alpha; beta; proto-oncogene; hematopoietic cell line;
KW   apoptosis; membrane-associated cytoplasmic protein; B cell; T cell;
KW   proliferation; cell cycle progression; Bax; apoptotic cell death;
KW   apoptosis; cytokine;  death repressor; BH1; BH2; cancer therapy;
KW   hyperplasia;  immunodeficiency disease; AIDS; neurodegeneration;
KW   ischaemic cell death.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   Domain          136..155
FT   /label= BH1_domain
FT   /note= "Represents Bax binding site"
FT   Domain          187..192
FT   /label= BH2_domain
FT   /note= "Represents Bax binding site"
PN   WO9505750-A.
PD   02-MAR-1995.
PF   24-AUG-1994; U09701.
PR   26-AUG-1993; US-112208.
PR   25-MAY-1994; US-248819.
PA   (UNIW ) UNIV WASHINGTON.
PI   Korsmeyer SJ;
DR   WPI; 95-106605/14.
PT   Methods for producing and identifying mutant bcl-2 proteins -
PT   that lack death repressor activity and/or lacks binding to Bax.
PS   Disclosure; Page 40; 133pp; English.
CC   The sequences given in R71404-05 represent the human bcl-2 alpha and
CC   beta proteins respectively.  bcl-2 is encoded by a proto-oncogene and
CC   is capable of inhibiting apoptosis in many hematopoietic cell systems.
CC   bcl-2 is a 26 kD membrane-associated cytoplasmic protein and is thought
CC   to function by enhancing the survival of hematopoietic cells of B and T
CC   origins rather than directly promoting proliferation of these cell
CC   types.  bcl-2 has not been shown to directly promote cell cycle
CC   progression nor does it necessarily alter the dose reponse to limiting
CC   concentrations of IL-3.  bcl-2 has been shown to form heterodimers with
CC   a 21 kD protein, Bax.  Overexpressed Bax accelerates apoptotic cell death
CC   induced by cytokine deprivation in an IL-3 dependent cell line, and it
CC   also acts to counter the death repressor activity of bcl-2.  Therefore,
CC   the ratio between bcl-2 and Bax determines cell survival or death
CC   following an apoptotic stimulus.  The invention gives a mutant form of
CC   bcl-2 in which there is at least one amino acid substitution or deletion
CC   in the BH1 or BH2 domains.  This makes the mutant protein substantially
CC   incapable of binding Bax and/or incapable of death repressor activity.
CC   Down regulation of bcl-2 is useful in cancer therapy, controlling
CC   hyperplasias and eliminating self-reactive clones in autoimmunity by
CC   favouring death effector molecules.  Up regulating bcl-2is beneficial in
CC   treatment and diagnosis of immunodeficiency diseases, including AIDS and
CC   neurodegenerative and ischaemic cell death.
SQ   Sequence   205 AA;

  Query Match           16.6%;  Score 259;  DB 13;  Length 205;
  Best Local Similarity  29.7%;  Pred. No. 3.72e-13;
```

```
  Matches    38;  Conservative   37;  Mismatches  49;  Indels   4;  Gaps   3;

Db      74 tpaapgaaagpalspvppvvhla--lrqagddfsrryrgdfaemssqlhltpftargrfa 131
           :|| |   : | | | : : :::  |   |||::||| ::|  |  :|: |:  |   |:
Qy      54 APADPEMVTLP-LQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENAYEYFT 112

Db     132 tvveelfrdgvnwgrivaffefggvmcvesvnremsplvdnialwmteylnrh-lhtwiq 190
            :   ||  |:||||:||:: ||  : :    : :: ::  ::  : ::: :| :  ||
Qy     113 KIATSLFESGINWGRVVALLGFGYRLALHVYQHGLTGFLGQVTRFVVDFMLHHCIARWIA 172

Db     191 dnggwvga 198
           : ||||:|
Qy     173 QRGGWVAA 180


Search completed: Wed Aug 20 11:00:33 1997
Job time : 45 secs.
```

****

MPSRCH (TM)

****

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Wed Aug 20 11:00:51 1997; MasPar time 8.88 Seconds
686.301 Million cell updates/sec

Tabular output not generated.

Title: >US-08-320-157-7
Description: (1-211) from US08320157.pep
Perfect Score: 1561
Sequence: 1 MASGQGPGPPRQECGEPALP.....LVVLGVVLLGQFVVRRFFKS 211

Scoring table: PAM 150
Gap 11

Searched: 91006 seqs, 28888923 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: pir51
1:ann1 2:ann2 3:ann3 4:ann4 5:unann1 6:unann2 7:unann3
8:unann4 9:unann5 10:unann6 11:unann7 12:unann8
13:unann9 14:unann10 15:unenc 16:unrev

Statistics: Mean 45.149; Variance 108.206; scale 0.417

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description	Pred. No.
1	1561	100.0	211	13	S58873	Bak protein - human	4.94e-244
2	1529	98.0	211	13	S58875	cdn-2 protein - huma	2.15e-238
3	274	17.6	190	13	A47537	apoptosis regulator	3.11e-24
4	270	17.3	233	14	S51761	BCL-X protein - rat	1.27e-23
5	270	17.3	233	14	I49056	bcl-x long - mouse	1.27e-23
6	266	17.0	233	13	B47537	apoptosis regulator	5.19e-23
7	263	16.8	214	14	I49057	bcl-x transmembrane	1.49e-22
8	262	16.8	216	6	B37332	transforming protein	2.11e-22
9	261	16.7	233	6	A37332	transforming protein	2.99e-22
10	259	16.6	205	2	TVHUB1	transforming protein	6.01e-22
11	255	16.3	206	6	D37332	transforming protein	2.43e-21
12	255	16.3	239	2	TVHUA1	transforming protein	2.43e-21
13	252	16.1	233	14	I67431	BCL-X-Long - rat	6.88e-21
14	250	16.0	199	2	TVMSB1	transforming protein	1.38e-20
15	250	16.0	236	2	TVMSA1	transforming protein	1.38e-20
16	250	16.0	237	6	E37332	transforming protein	1.38e-20
17	247	15.8	236	14	I67432	BCL-2 - rat (fragmen	3.89e-20
18	239	15.3	232	6	S24390	transforming protein	6.14e-19
19	217	13.9	192	14	D47538	programmed cell deat	1.09e-15
20	206	13.2	192	13	A47538	bcl-2-associated pro	4.28e-14
21	201	12.9	218	13	B47538	bcl-2-associated pro	2.24e-13
22	192	12.3	350	13	A47476	BCL2 homolog MCL1 -	4.28e-12
23	186	11.9	133	14	I53295	bax - rat (fragment)	3.00e-11
24	183	11.7	154	14	I58194	gene bcl-2 protein -	7.87e-11
25	177	11.3	143	13	I38921	BAX splice form delt	5.36e-10
26	176	11.3	172	14	I49449	hemopoietic-specific	7.37e-10
27	114	7.3	890	8	S44150	coat protein - straw	6.55e-02
28	112	7.2	255	5	S26032	cytochrome-c oxidase	1.11e-01
29	111	7.1	177	13	S54778	NR-13 protein - quai	1.44e-01
30	107	6.9	400	10	S35958	beta-glucosidase - S	4.07e-01
31	107	6.9	470	2	SYPSRA	threonine synthase (	4.07e-01
32	106	6.8	133	4	GPBPP4	gop protein - satell	5.25e-01
33	100	6.4	560	7	S46724	hexose transport pro	2.35e+00
34	100	6.4	567	7	S31294	hexose transport pro	2.35e+00
35	100	6.4	567	11	S70110	HXT3 protein - yeast	2.35e+00
36	97	6.2	170	14	I49055	bcl-x short - mouse	4.88e+00
37	97	6.2	814	1	CZCLEM	cellulase (EC 3.2.1.	4.88e+00
38	95	6.1	206	11	B28443	phosphatidyl-N-methy	7.87e+00
39	95	6.1	471	7	A39024	collagen alpha 3(IV)	7.87e+00
40	96	6.1	527	16	S62484	hypothetical protein	6.20e+00
41	95	6.1	552	16	S65787	glutamyl-tRNA synthe	7.87e+00
42	95	6.1	569	7	S50771	sugar transport prot	7.87e+00
43	95	6.1	817	1	RRVGCT	RNA-directed RNA pol	7.87e+00
44	95	6.1	885	11	S22389	acetylglutamate kina	7.87e+00
45	96	6.1	2007	3	B43402	myosin heavy chain-B	6.20e+00

ALIGNMENTS

RESULT 1
ENTRY S58873 #type complete
TITLE Bak protein - human
ALTERNATE_NAMES bcl-2 homolog; cdn-1 protein
ORGANISM #formal_name Homo sapiens #common_name man
DATE 15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 06-Sep-1996
ACCESSIONS S58873; S58872; S58874
REFERENCE S58873
#authors Chittenden, T.; Harrington, E.A.; O'Connor, R.; Flemington, C.; Lutz, R.J.; Evan, G.I.; Guild, B.C.
#journal Nature (1995) 374:733-736
#title Induction of apoptosis by the Bcl-2 homologue Bak.
#accession S58873
##status preliminary; nucleic acid sequence not shown
##molecule_type mRNA
##residues 1-211 ##label CHI
##cross-references EMBL:U23765
REFERENCE S58872
#authors Farrow, S.N.; White, J.H.M.; Martinou, I.; Raven, T.; Pun, K.T.; Grinham, C.J.; Martinou, J.C.; Brown, R.
#journal Nature (1995) 374:731-733
#title Cloning of a bcl-2 homologue by interaction with adenovirus E1B 19K.
#accession S58872
##status preliminary
##molecule_type mRNA
##residues 1-211 ##label FAR
##cross-references EMBL:X84213
REFERENCE S58874
#authors Kiefer, M.C.; Brauer, M.J.; Powers, V.C.; Wu, J.J.; Umansky, S.R.; Tomei, L.D.; Barr, P.J.
#journal Nature (1995) 374:736-739
#title Modulation of apoptosis by the widely distributed Bcl-2 homologue Bak.
#accession S58874
##status preliminary
##molecule_type mRNA
##residues 1-211 ##label KIE
##cross-references EMBL:U16811
GENETICS
#gene GDB:BAK-LSB
##cross-references GDB:635887
SUMMARY #length 211 #molecular-weight 23409 #checksum 801

```
  Query Match          100.0%;  Score 1561;  DB 13;  Length 211;
  Best Local Similarity 100.0%;  Pred. No. 4.94e-244;
  Matches   211;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      1 masgqgpgpprqecgepalpsaseeqvaqdteevfrsyvfyrhqqeqeaegvaapadpem 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      1 MASGQGPGPPRQECGEPALPSASEEQVAQDTEEVFRSYVFYRHQQEQEAEGVAAPADPEM 60

Db     61 vtlplqpsstmgqvgrqlaiigddinrrydsefqtmlqhlqptaenayeyftkiatslfe 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     61 VTLPLQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENAYEYFTKIATSLFE 120

Db    121 sginwgrvvallgfgyrlalhvyqhgltgflgqvtrfvvdfmlhhciarwiaqrggwvaa 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    121 SGINWGRVVALLGFGYRLALHVYQHGLTGFLGQVTRFVVDFMLHHCIARWIAQRGGWVAA 180

Db    181 lnlgngpilnvlvvlgvvllgqfvvrrffks 211
          |||||||||||||||||||||||||||||||
Qy    181 LNLGNGPILNVLVVLGVVLLGQFVVRRFFKS 211


RESULT    2
ENTRY             S58875     #type complete
TITLE             cdn-2 protein - human
ORGANISM          #formal_name Homo sapiens #common_name man
DATE              15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change
                    01-Mar-1996
ACCESSIONS        S58875
REFERENCE         S58874
   #authors       Kiefer, M.C.; Brauer, M.J.; Powers, V.C.; Wu, J.J.; Umansky,
                    S.R.; Tomei, L.D.; Barr, P.J.
   #journal       Nature (1995) 374:736-739
   #title         Modulation of apoptosis by the widely distributed Bcl-2
                    homologue Bak.
   #accession     S58875
      ##status         preliminary; nucleic acid sequence not shown;
                         translation not shown
      ##molecule_type DNA
      ##residues       1-211 ##label KIE
      ##cross-references EMBL:U16812
      ##note           the nucleotide sequence was submitted to the EMBL Data
                         Library, November 1994
SUMMARY           #length 211  #molecular-weight 23411  #checksum 9485

  Query Match           98.0%;  Score 1529;  DB 13;  Length 211;
  Best Local Similarity 97.2%;  Pred. No. 2.15e-238;
  Matches   205;  Conservative    5;  Mismatches   1;  Indels   0;  Gaps   0;

Db      1 masgqgpgpprqecgepalpsaseeqvaqdteevfrsyvfyhhqqeqeaegaaapadpem 60
          ||||||||||||||||||||||||||||||||||||||:||||||||| ||||||||
Qy      1 MASGQGPGPPRQECGEPALPSASEEQVAQDTEEVFRSYVFYRHQQEQEAEGVAAPADPEM 60

Db     61 vtlplqpsstmgqvgrqlaiigddinrrydsefqtmlqhlqptaenayeyftkiasslfe 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||
Qy     61 VTLPLQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENAYEYFTKIATSLFE 120

Db    121 sginwgrvvallgfsyrlalhiyqrgltgflgqvtrfvvdfmlhhciarwiaqrggwvaa 180
          |||||||||||||:||||||:||:||||||||||||||||||||||||||||||||||||
Qy    121 SGINWGRVVALLGFGYRLALHVYQHGLTGFLGQVTRFVVDFMLHHCIARWIAQRGGWVAA 180

Db    181 lnlgngpilnvlvvlgvvllgqfvvrrffks 211
          |||||||||||||||||||||||||||||||
Qy    181 LNLGNGPILNVLVVLGVVLLGQFVVRRFFKS 211


RESULT    3
ENTRY             A47537     #type complete
TITLE             apoptosis regulator bcl-x - chicken
ORGANISM          #formal_name Gallus gallus #common_name chicken
DATE              03-May-1994 #sequence_revision 03-May-1994 #text_change
                    03-May-1994
ACCESSIONS        A47537
REFERENCE         A47537
   #authors       Boise, L.H.; Gonzalez-Garcia, M.; Postema, C.E.; Ding, L.;
                    Lindsten, T.; Turka, L.A.; Mao, X.; Nunez, G.; Thompson,
                    C.B.
   #journal       Cell (1993) 74:597-608
   #title         bcl-x, a bcl-2-related gene that functions as a dominant
                    regulator of apoptotic cell death.
   #accession     A47537
      ##status         preliminary
      ##molecule_type DNA
      ##residues       1-190 ##label BOI
      ##cross-references GB:L20120
SUMMARY           #length 190  #molecular-weight 21467  #checksum 5509

  Query Match           17.6%;  Score 274;  DB 13;  Length 190;
  Best Local Similarity 28.2%;  Pred. No. 3.11e-24;
  Matches    37;  Conservative   33;  Mismatches  59;  Indels   2;  Gaps   2;

Db     60 vvngatvhrsslevheivrasdvrqalrdagdefelryrrafsdltsqlhitpgtayqsf 119
          |   |     :| ::     ::| : |   ||::: ||   |  :  :|: |:  ||: |
Qy     52 VAAPADPEMVTLPLQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENAYEYF 111

Db    120 eqvvnelfhdgvnwgrivaffsfggalcvesvdkemrvlvgrivswmttyltdh-ldpwi 178
            :   ||  |:||||:||:::||  | :   :  :  ::|::   :  ::  | :  ||
Qy    112 TKIATSLFESGINWGRVVALLGFGYRLALHVYQHGLTGFLGQVTRFVVDFMLHHCIARWI 171

Db    179 qenggwvrtal 189
          : |||| :||
Qy    172 AQRGGWV-AAL 181


RESULT    4
ENTRY             S51761     #type complete
TITLE             BCL-X protein - rat
ORGANISM          #formal_name Rattus norvegicus #common_name Norway rat
DATE              07-May-1995 #sequence_revision 01-Sep-1995 #text_change
                    01-Dec-1995
ACCESSIONS        S51761; S51762
REFERENCE         S51761
   #authors       Michaelidis, T.M.
   #submission    submitted to the EMBL Data Library, November 1994
   #accession     S51761
      ##status         preliminary
      ##molecule_type DNA
      ##residues       1-233 ##label MIC
      ##cross-references EMBL:X82537
REFERENCE         S51761
   #authors       Michaelidis, T.M.
   #submission    submitted to the EMBL Data Library, November 1994
   #accession     S51762
      ##status         preliminary
      ##molecule_type DNA
      ##residues       1-125,189-233 ##label MI2
      ##cross-references EMBL:X82537
GENETICS
   #introns       125/3
SUMMARY           #length 233  #molecular-weight 26130  #checksum 6378

  Query Match           17.3%;  Score 270;  DB 14;  Length 233;
  Best Local Similarity 25.5%;  Pred. No. 1.27e-23;
  Matches    36;  Conservative   42;  Mismatches  60;  Indels   3;  Gaps   3;

Db     61 dspavngatghs-ssldarevipmaavkqalreagdefelryrrafsdltsqlhitpgta 119
          :: :| :::     :|  :    |: | : |   ||::: ||   |  :  :|: |:  |
Qy     48 EAEGVAAPADPEMVTLPLQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENA 107

Db    120 yqsfeqvvnelfrdgvnwgrivaffsfggalcvesvdkemqvlvsriaswmatylndh-l 178
          |: |  :   ||  |:||||:||:::||  | :   :  :  ::::::  :  ::  | :
Qy    108 YEYFTKIATSLFESGINWGRVVALLGFGYRLALHVYQHGLTGFLGQVTRFVVDFMLHHCI 167

Db    179 epwiqenggwdtfvdlygnna 199
            || : ||| : ::| || :
Qy    168 ARWIAQRGGWVAALNL-GNGP 187
```

```
RESULT    5
ENTRY             I49056     #type complete
TITLE             bcl-x long - mouse
ORGANISM          #formal_name Mus musculus #common_name house mouse
DATE              02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change
                    15-Oct-1996
ACCESSIONS        I49056; S52866
REFERENCE         I49055
   #authors       Fang, W.; Rivard, J.J.; Mueller, D.L.; Behrens, T.W.
   #journal       J. Immunol. (1994) 153:4388-4398
   #title         Cloning and molecular characterization of mouse bcl-x in B
                    and T lymphocytes.
   #cross-references MUID:95052604
   #accession     I49056
      ##status         preliminary; translated from GB/EMBL/DDBJ
      ##molecule_type mRNA
      ##residues       1-233 ##label RES
      ##cross-references EMBL:U10101; NID:g506647; CDS_PID:g506648
REFERENCE         S52866
   #authors       Kamesaki, H.; Michaud, G.Y.; Takatsu, K.; Okuma, M.
   #submission    submitted to the EMBL Data Library, November 1994
   #description   IL-5 inhibits anti-IgM-induced apoptosis in an immature B
                    cell line through inductin of bcl-Xl.
   #accession     S52866
      ##status         preliminary
      ##molecule_type mRNA
      ##residues       1-233 ##label KAM
      ##cross-references EMBL:X83574
SUMMARY           #length 233  #molecular-weight 26132  #checksum 5739

  Query Match            17.3%;  Score 270;  DB 14;  Length 233;
  Best Local Similarity  25.5%;  Pred. No. 1.27e-23;
  Matches    36;  Conservative   42;  Mismatches  60;  Indels   3;  Gaps   3;

Db      61 dspavngatghs-ssldarevipmaavkqalreagdefelryrrafsdltsqlhitpgta 119
           :: :| :::     :|  :    |: | : |   ||::: ||   |  :  :|: |:  |
Qy      48 EAEGVAAPADPEMVTLPLQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENA 107

Db     120 yqsfeqvvnelfrdgvnwgrivaffsfggalcvesvdkemqvlvsriaswmatylndh-l 178
           |: |  :   ||  |:||||:||:::||  | :   :  :  ::::::  :  ::  | :
Qy     108 YEYFTKIATSLFESGINWGRVVALLGFGYRLALHVYQHGLTGFLGQVTRFVVDFMLHHCI 167

Db     179 epwiqenggwdtfvdlygnna 199
             || : ||| : ::| || :
Qy     168 ARWIAQRGGWVAALNL-GNGP 187


RESULT    6
ENTRY             B47537     #type complete
TITLE             apoptosis regulator bcl-xL - human
ALTERNATE_NAMES   bcl-2-related protein
CONTAINS          apoptosis regulator bcl-xS
ORGANISM          #formal_name Homo sapiens #common_name man
DATE              16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change
                    06-Sep-1996
ACCESSIONS        B47537; C47537
REFERENCE         A47537
   #authors       Boise, L.H.; Gonzalez-Garcia, M.; Postema, C.E.; Ding, L.;
                    Lindsten, T.; Turka, L.A.; Mao, X.; Nunez, G.; Thompson,
                    C.B.
   #journal       Cell (1993) 74:597-608
   #title         bcl-x, a bcl-2-related gene that functions as a dominant
                    regulator of apoptotic cell death.
   #accession     B47537
      ##status         nucleic acid sequence not shown; translated from
                         GB/EMBL/DDBJ
      ##molecule_type mRNA
      ##residues       1-233 ##label BOI
      ##cross-references GB:L20121; CDS_PID:Q07817
   #accession     C47537
      ##status         nucleic acid sequence not shown; translated from
                         GB/EMBL/DDBJ
      ##molecule_type mRNA
      ##residues       1-69,'G',71-125,189-233 ##label BO2
      ##cross-references GB:L20122; CDS_PID:g623237
GENETICS
   #gene          GDB:BCL2L
      ##cross-references GDB:228079
KEYWORDS          alternative splicing; apoptosis
FEATURE
   1-233               #product apoptosis regulator bcl-xL #status predicted
                         #label MAT\
   1-125,189-233       #product apoptosis regulator bcl-xS #status predicted
                         #label MA2
SUMMARY           #length 233  #molecular-weight 26063  #checksum 5340

  Query Match            17.0%;  Score 266;  DB 13;  Length 233;
  Best Local Similarity  25.5%;  Pred. No. 5.19e-23;
  Matches    36;  Conservative   42;  Mismatches  60;  Indels   3;  Gaps   3;

Db      61 dspavngatahs-ssldarevipmaavkqalreagdefelryrrafsdltsqlhitpgta 119
           :: :| :::     :|  :    |: | : |   ||::: ||   |  :  :|: |:  |
Qy      48 EAEGVAAPADPEMVTLPLQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENA 107

Db     120 yqsfeqvvnelfrdgvnwgrivaffsfggalcvesvdkemqvlvsriaawmatylndh-l 178
           |: |  :   ||  |:||||:||:::||  | :   :  :  ::::::  :  ::  | :
Qy     108 YEYFTKIATSLFESGINWGRVVALLGFGYRLALHVYQHGLTGFLGQVTRFVVDFMLHHCI 167

Db     179 epwiqenggwdtfvelygnna 199
             || : ||| : ::| || :
Qy     168 ARWIAQRGGWVAALNL-GNGP 187


RESULT    7
ENTRY             I49057     #type complete
TITLE             bcl-x transmembrane deleted - mouse
ORGANISM          #formal_name Mus musculus #common_name house mouse
DATE              02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change
                    02-Jul-1996
ACCESSIONS        I49057
REFERENCE         I49055
   #authors       Fang, W.; Rivard, J.J.; Mueller, D.L.; Behrens, T.W.
   #journal       J. Immunol. (1994) 153:4388-4398
   #title         Cloning and molecular characterization of mouse bcl-x in B
                    and T lymphocytes.
   #cross-references MUID:95052604
   #accession     I49057
      ##status         preliminary; translated from GB/EMBL/DDBJ
      ##molecule_type mRNA
      ##residues       1-214 ##label RES
      ##cross-references EMBL:U10102; NID:g506649; CDS_PID:g506650
GENETICS
   #note          gene name bcl-x long
SUMMARY           #length 214  #molecular-weight 23900  #checksum 9730

  Query Match            16.8%;  Score 263;  DB 14;  Length 214;
  Best Local Similarity  25.4%;  Pred. No. 1.49e-22;
  Matches    33;  Conservative   38;  Mismatches  57;  Indels   2;  Gaps   2;

Db      61 dspavngatghs-ssldarevipmaavkqalreagdefelryrrafsdltsqlhitpgta 119
           :: :| :::     :|  :    |: | : |   ||::: ||   |  :  :|: |:  |
Qy      48 EAEGVAAPADPEMVTLPLQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENA 107

Db     120 yqsfeqvvnelfrdgvnwgrivaffsfggalcvesvdkemqvlvsriaswmatylndh-l 178
           |: |  :   ||  |:||||:||:::||  | :   :  :  ::::::  :  ::  | :
Qy     108 YEYFTKIATSLFESGINWGRVVALLGFGYRLALHVYQHGLTGFLGQVTRFVVDFMLHHCI 167

Db     179 epwiqenggw 188
             || : |||
Qy     168 ARWIAQRGGW 177


RESULT    8
ENTRY             B37332     #type complete
TITLE             transforming protein (bcl-2-beta) - chicken
```

```
ORGANISM          #formal_name Gallus gallus #common_name chicken
DATE              03-Mar-1993 #sequence_revision 03-Mar-1993 #text_change
                    15-Jun-1996
ACCESSIONS        B37332
REFERENCE         A37332
   #authors       Eguchi, Y.; Ewert, D.L.; Tsujimoto, Y.
   #journal       Nucleic Acids Res. (1992) 20:4187-4192
   #title         Isolation and characterization of the chicken bcl-2 gene:
                    expression in a variety of tissues including lymphoid and
                    neuronal organs in adult and embryo.
   #accession     B37332
      ##status         preliminary; nucleic acid sequence not shown; not
                         compared with conceptual translation
      ##molecule_type DNA
      ##residues      1-216 ##label EGU
      ##cross-references EMBL:D11382
CLASSIFICATION    #superfamily bcl transforming protein
SUMMARY           #length 216  #molecular-weight 23492  #checksum 3596

  Query Match             16.8%;  Score 262;  DB 6;  Length 216;
  Best Local Similarity   28.5%;  Pred. No. 2.11e-22;
  Matches    39;  Conservative   38;  Mismatches  56;  Indels   4;  Gaps   4;

Db      60 hhrpeppgsaaasevppae-glrpapp-g-vhlalrqagdefsrryqrdfaqmsgqlhlt 116
           ::  |: | ||:::       |:|::  | |    |   ||:::|||: :|  |  :|: |
Qy      44 QQEQEAEGVAAPADPEMVTLPLQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPT 103

Db     117 pftahgrfvavveelfrdgvnwgrivaffefggvmcvesvnremsplvdniatwmteyln 176
           :  |   |  :   ||  |:||||:||:: ||  : :    : :: ::  ::  : :::
Qy     104 AENAYEYFTKIATSLFESGINWGRVVALLGFGYRLALHVYQHGLTGFLGQVTRFVVDFML 163

Db     177 rh-lhnwiqdnggwvra 192
           :| :  || : |||| |
Qy     164 HHCIARWIAQRGGWVAA 180


RESULT    9
ENTRY             A37332     #type complete
TITLE             transforming protein (bcl-2-alpha) - chicken
ORGANISM          #formal_name Gallus gallus #common_name chicken
DATE              03-Mar-1993 #sequence_revision 03-Mar-1993 #text_change
                    12-Apr-1995
ACCESSIONS        A37332; S35453
REFERENCE         A37332
   #authors       Eguchi, Y.; Ewert, D.L.; Tsujimoto, Y.
   #journal       Nucleic Acids Res. (1992) 20:4187-4192
   #title         Isolation and characterization of the chicken bcl-2 gene:
                    expression in a variety of tissues including lymphoid and
                    neuronal organs in adult and embryo.
   #accession     A37332
      ##status         nucleic acid sequence not shown
      ##molecule_type DNA
      ##residues      1-233 ##label EGU
      ##cross-references EMBL:D11381
GENETICS
   #introns       189/3
CLASSIFICATION    #superfamily bcl transforming protein
KEYWORDS          transforming protein
SUMMARY           #length 233  #molecular-weight 25687  #checksum 99

  Query Match             16.7%;  Score 261;  DB 6;  Length 233;
  Best Local Similarity   28.5%;  Pred. No. 2.99e-22;
  Matches    41;  Conservative   41;  Mismatches  57;  Indels   5;  Gaps   5;

Db      60 hhrpeppgsaaasevppae-glrpapp-g-vhlalrqagdefsrryqrdfaqmsgqlhlt 116
           ::  |: | ||:::       |:|::  | |    |   ||:::|||: :|  |  :|: |
Qy      44 QQEQEAEGVAAPADPEMVTLPLQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPT 103

Db     117 pftahgrfvavveelfrdgvnwgrivaffefggvmcvesvnremsplvdniatwmteyln 176
           :  |   |  :   ||  |:||||:||:: ||  : :    : :: ::  ::  : :::
Qy     104 AENAYEYFTKIATSLFESGINWGRVVALLGFGYRLALHVYQHGLTGFLGQVTRFVVDFML 163

Db     177 rh-lhnwiqdnggwdafvelygns 199
           :| :  || : |||| | ::| ||:
Qy     164 HHCIARWIAQRGGWVAALNL-GNG 186


RESULT   10
ENTRY             TVHUB1     #type complete
TITLE             transforming protein bcl-2-beta - human
ORGANISM          #formal_name Homo sapiens #common_name man
DATE              31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change
                    06-Sep-1996
ACCESSIONS        B29409
REFERENCE         A29409
   #authors       Tsujimoto, Y.; Croce, C.M.
   #journal       Proc. Natl. Acad. Sci. U.S.A. (1986) 83:5214-5218
   #title         Analysis of the structure, transcripts, and protein products
                    of bcl-2, the gene involved in human follicular lymphoma.
   #cross-references MUID:86259760
   #accession     B29409
      ##molecule_type mRNA
      ##residues      1-205 ##label TSU
GENETICS
   #gene          GDB:BCL2
      ##cross-references GDB:119031
   #map_position 18q21.33-18q21.33
CLASSIFICATION    #superfamily bcl transforming protein
KEYWORDS          alternative splicing; B-cell lymphoma; follicular lymphoma;
                    transforming protein
SUMMARY           #length 205  #molecular-weight 22182  #checksum 1183

  Query Match             16.6%;  Score 259;  DB 2;  Length 205;
  Best Local Similarity   29.7%;  Pred. No. 6.01e-22;
  Matches    38;  Conservative   37;  Mismatches  49;  Indels   4;  Gaps   3;

Db      74 tpaapgaaagpalspvppvvhla--lrqagddfsrryrgdfaemssqlhltpftargrfa 131
           :|| |   : | | | : : :::  |   |||::||| ::|  |  :|: |:  |   |:
Qy      54 APADPEMVTLP-LQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENAYEYFT 112

Db     132 tvveelfrdgvnwgrivaffefggvmcvesvnremsplvdnialwmteylnrh-lhtwiq 190
           :   ||  |:||||:||:: ||  : :    : :: ::  ::  : ::: :| :  ||
Qy     113 KIATSLFESGINWGRVVALLGFGYRLALHVYQHGLTGFLGQVTRFVVDFMLHHCIARWIA 172

Db     191 dnggwvga 198
           : ||||:|
Qy     173 QRGGWVAA 180


RESULT   11
ENTRY             D37332     #type complete
TITLE             transforming protein (bcl-2-beta) - human
ORGANISM          #formal_name Homo sapiens #common_name man
DATE              03-Mar-1993 #sequence_revision 03-Mar-1993 #text_change
                    15-Jun-1996
ACCESSIONS        D37332
REFERENCE         A37332
   #authors       Eguchi, Y.; Ewert, D.L.; Tsujimoto, Y.
   #journal       Nucleic Acids Res. (1992) 20:4187-4192
   #title         Isolation and characterization of the chicken bcl-2 gene:
                    expression in a variety of tissues including lymphoid and
                    neuronal organs in adult and embryo.
   #accession     D37332
      ##status         preliminary; nucleic acid sequence not shown; not
                         compared with conceptual translation
      ##molecule_type DNA
      ##residues      1-206 ##label EGU
CLASSIFICATION    #superfamily bcl transforming protein
SUMMARY           #length 206  #molecular-weight 22440  #checksum 5581

  Query Match             16.3%;  Score 255;  DB 6;  Length 206;
  Best Local Similarity   29.7%;  Pred. No. 2.43e-21;
  Matches    38;  Conservative   35;  Mismatches  51;  Indels   4;  Gaps   3;

Db      75 tpaapgaaagpalspvppvvhlt--lrqagddfsrryrrdfaemssqlhltpftargrfa 132
           :|| |   : | | | : : ::   |   |||::|||  :|  |  :|: |:  |   |:
```

```
Qy     54 APADPEMVTLP-LQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENAYEYFT 112

Db    133 tvveelfrdgvnwgrivaffefggvmcvesvnremsplvdnialwmteylnrh-lhtwiq 191
          :   ||  |:||||:||:: ||  : :    : :: ::  ::  : ::: :| :  ||
Qy    113 KIATSLFESGINWGRVVALLGFGYRLALHVYQHGLTGFLGQVTRFVVDFMLHHCIARWIA 172

Db    192 dnggwvga 199
          : ||||:|
Qy    173 QRGGWVAA 180


RESULT   12
ENTRY            TVHUA1     #type complete
TITLE            transforming protein bcl-2, splice form alpha - human
ORGANISM         #formal_name Homo sapiens #common_name man
DATE             31-Dec-1988 #sequence_revision 07-Jun-1996 #text_change
                   06-Sep-1996
ACCESSIONS       C37332; A29409; S02452; A24428; A27622; B27622
REFERENCE        A37332
   #authors      Eguchi, Y.; Ewert, D.L.; Tsujimoto, Y.
   #journal      Nucleic Acids Res. (1992) 20:4187-4192
   #title        Isolation and characterization of the chicken bcl-2 gene:
                   expression in a variety of tissues including lymphoid and
                   neuronal organs in adult and embryo.
   #accession    C37332
      ##status        nucleic acid sequence not shown; not compared with
                        conceptual translation
      ##molecule_type DNA
      ##residues      1-239 ##label EGU
      ##note          this report is a correction
REFERENCE        A29409
   #authors      Tsujimoto, Y.; Croce, C.M.
   #journal      Proc. Natl. Acad. Sci. U.S.A. (1986) 83:5214-5218
   #title        Analysis of the structure, transcripts, and protein products
                   of bcl-2, the gene involved in human follicular lymphoma.
   #cross-references MUID:86259760
   #accession    A29409
      ##molecule_type mRNA
      ##residues      1-95,'A',97-109,'G',111-236,'S',238-239 ##label TSU
      ##note          this sequence has been corrected in reference A37332
REFERENCE        S02452
   #authors      Seto, M.; Jaeger, U.; Hockett, R.D.; Graninger, W.; Bennett,
                   S.; Goldman, P.; Korsmeyer, S.J.
   #journal      EMBO J. (1988) 7:123-131
   #title        Alternative promoters and exons, somatic mutation and
                   deregulation of the Bcl-2--Ig fusion gene in lymphoma.
   #cross-references MUID:88196071
   #accession    S02452
      ##molecule_type mRNA
      ##residues      1-239 ##label SET
REFERENCE        A24428
   #authors      Cleary, M.L.; Smith, S.D.; Sklar, J.
   #journal      Cell (1986) 47:19-28
   #title        Cloning and structural analysis of cDNAs for bcl-2 and a
                   hybrid bcl-2/immunoglobulin transcript resulting from the t
                   (14;18) translocation.
   #cross-references MUID:87002488
   #accession    A24428
      ##molecule_type mRNA
      ##residues      1-58,'T',60-116,'R',118-239 ##label CLE
REFERENCE        A27622
   #authors      Hua, C.; Zorn, S.; Jensen, J.P.; Coupland, R.W.; Ko, H.S.;
                   Wright, J.J.; Bakhshi, A.
   #journal      Oncogene Res. (1988) 2:263-275
   #title        Consequences of the t(14;18) chromosomal translocation in
                   follicular lymphoma: deregulated expression of a chimeric
                   and mutated BCL-2 gene.
   #cross-references MUID:88217344
   #accession    A27622
      ##molecule_type mRNA
      ##residues      1-58,'T',60-239 ##label HUA
   #accession    B27622
      ##molecule_type DNA
      ##residues      1-6,'S',8-58,'T',60-128,'C',130-239 ##label HUA2
      ##note          the sequence was determined from the germline gene
COMMENT          Constitutive expression of BCL2 following t(14:18) chromosomal
                   translocation is typically found in follicular lymphoma.
GENETICS
   #gene         GDB:BCL2
      ##cross-references GDB:119031
   #map_position 18q21.33-18q21.33
FUNCTION
   #description  blocks apoptosis in hematopoietic cells
CLASSIFICATION   #superfamily bcl transforming protein
KEYWORDS         alternative splicing; B-cell lymphoma; follicular lymphoma;
                   proto-oncogene; transforming protein
SUMMARY          #length 239  #molecular-weight 26266  #checksum 8323

  Query Match           16.3%;  Score 255;  DB 2;  Length 239;
  Best Local Similarity  29.3%;  Pred. No. 2.43e-21;
  Matches    41;  Conservative   37;  Mismatches  57;  Indels   5;  Gaps   4;

Db     74 tpaapgaaagpalspvppvvhlt--lrqagddfsrryrrdfaemssqlhltpftargrfa 131
          :|| |   : | | | : : ::   |   |||::|||  :|  |  :|: |:  |   |:
Qy     54 APADPEMVTLP-LQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENAYEYFT 112

Db    132 tvveelfrdgvnwgrivaffefggvmcvesvnremsplvdnialwmteylnrh-lhtwiq 190
          :   ||  |:||||:||:: ||  : :    : :: ::  ::  : ::: :| :  ||
Qy    113 KIATSLFESGINWGRVVALLGFGYRLALHVYQHGLTGFLGQVTRFVVDFMLHHCIARWIA 172

Db    191 dnggwdafvel-ygpsmrpl 209
          : ||| | ::|  || :  |
Qy    173 QRGGWVAALNLGNGPILNVL 192


RESULT   13
ENTRY            I67431     #type complete
TITLE            BCL-X-Long - rat
ORGANISM         #formal_name Rattus norvegicus #common_name Norway rat
DATE             26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change
                   26-Jul-1996
ACCESSIONS       I67431
REFERENCE        I53295
   #authors      Tilly, J.L.; Tilly, K.I.; Kenton, M.L.; Johnson, A.L.
   #journal      Endocrinology (1995) 136:232-241
   #title        Expression of members of the bcl-2 gene family in the
                   immature rat ovary: equine chorionic gonadotropin-mediated
                   inhibition of granulosa cell apoptosis is associated with
                   decreased bax and constitutive bcl-2 and bcl-xlong
                   messenger ribonucleic acid levels.
   #cross-references MUID:95129487
   #accession    I67431
      ##status        preliminary; translated from GB/EMBL/DDBJ
      ##molecule_type mRNA
      ##residues      1-233 ##label RES
      ##cross-references EMBL:U34963; NID:g1004376; CDS_PID:g1004377
SUMMARY          #length 233  #molecular-weight 26122  #checksum 8310

  Query Match           16.1%;  Score 252;  DB 14;  Length 233;
  Best Local Similarity  28.4%;  Pred. No. 6.88e-21;
  Matches    33;  Conservative   31;  Mismatches  50;  Indels   2;  Gaps   2;

Db     83 maavkqalreagdefelryrrafsdltsqlhitpgtvyqsfeqvvnelfrdgvnwgriva 142
          |: | : |   ||::: ||   |  :  :|: |:   |: | :   ||  |:||||:||
Qy     71 MGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENAYEYFTKIATSLFESGINWGRVVA 130

Db    143 sssfggalcvesvdkemqvlvsriaswmatylndh-lepwiqenggwdtfvdlygn 197
            :||  | :   :  : ::::::  :  ::  | :  || : ||| : ::| ||
Qy    131 LLGFGYRLALHVYQHGLTGFLGQVTRFVVDFMLHHCIARWIAQRGGWVAALNL-GN 185


RESULT   14
ENTRY            TVMSB1     #type complete
TITLE            transforming protein bcl-2-beta - mouse
ORGANISM         #formal_name Mus musculus #common_name house mouse
DATE             31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change
```

```
                  02-Jun-1994
ACCESSIONS        B25960
REFERENCE         A90893
   #authors       Negrini, M.; Silini, E.; Kozak, C.; Tsujimoto, Y.; Croce,
                    C.M.
   #journal       Cell (1987) 49:455-463
   #title         Molecular analysis of mbcl-2: structure and expression of the
                    murine gene homologous to the human gene involved in
                    follicular lymphoma.
   #cross-references MUID:87187643
   #accession     B25960
      ##molecule_type DNA
      ##residues      1-199 ##label NEG
GENETICS
   #gene          BCL2
CLASSIFICATION    #superfamily bcl transforming protein
KEYWORDS          alternative splicing; transforming protein
SUMMARY           #length 199  #molecular-weight 22299  #checksum 7397

  Query Match           16.0%;  Score 250;  DB 2;  Length 199;
  Best Local Similarity  31.1%;  Pred. No. 1.38e-20;
  Matches    32;  Conservative   29;  Mismatches  41;  Indels   1;  Gaps   1;

Db     94 lrragddfsrryrrdfaemssqlhltpftargrfatvveelfrdgvnwgrivaffefggv 153
          |   |||::|||  :|  |  :|: |:  |   |: :   ||  |:||||:||:: ||
Qy     78 LAIIGDDINRRYDSEFQTMLQHLQPTAENAYEYFTKIATSLFESGINWGRVVALLGFGYR 137

Db    154 mcvesvnremsplvdnialwmteylnrh-lhtwiqdnggwvga 195
          : :    : :: ::  ::  : ::: :| :  || : ||||:|
Qy    138 LALHVYQHGLTGFLGQVTRFVVDFMLHHCIARWIAQRGGWVAA 180


RESULT   15
ENTRY             TVMSA1     #type complete
TITLE             transforming protein bcl-2-alpha - mouse
ORGANISM          #formal_name Mus musculus #common_name house mouse
DATE              31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change
                    02-Jun-1994
ACCESSIONS        A25960
REFERENCE         A90893
   #authors       Negrini, M.; Silini, E.; Kozak, C.; Tsujimoto, Y.; Croce,
                    C.M.
   #journal       Cell (1987) 49:455-463
   #title         Molecular analysis of mbcl-2: structure and expression of the
                    murine gene homologous to the human gene involved in
                    follicular lymphoma.
   #cross-references MUID:87187643
   #accession     A25960
      ##molecule_type DNA
      ##residues      1-236 ##label NEG
GENETICS
   #gene          BCL2
   #introns       192/3
CLASSIFICATION    #superfamily bcl transforming protein
KEYWORDS          alternative splicing; transforming protein
SUMMARY           #length 236  #molecular-weight 26524  #checksum 6709

  Query Match           16.0%;  Score 250;  DB 2;  Length 236;
  Best Local Similarity  30.4%;  Pred. No. 1.38e-20;
  Matches    35;  Conservative   31;  Mismatches  47;  Indels   2;  Gaps   2;

Db     94 lrragddfsrryrrdfaemssqlhltpftargrfatvveelfrdgvnwgrivaffefggv 153
          |   |||::|||  :|  |  :|: |:  |   |: :   ||  |:||||:||:: ||
Qy     78 LAIIGDDINRRYDSEFQTMLQHLQPTAENAYEYFTKIATSLFESGINWGRVVALLGFGYR 137

Db    154 mcvesvnremsplvdnialwmteylnrh-lhtwiqdnggwdafvel-ygpsmrpl 206
          : :    : :: ::  ::  : ::: :| :  || : ||| | ::|  || :  |
Qy    138 LALHVYQHGLTGFLGQVTRFVVDFMLHHCIARWIAQRGGWVAALNLGNGPILNVL 192


Search completed: Wed Aug 20 11:01:37 1997
Job time : 46 secs.
```

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Wed Aug 20 11:01:54 1997; MasPar time 6.20 Seconds
722.269 Million cell updates/sec

Tabular output not generated.

Title: >US-08-320-157-7
Description: (1-211) from US08320157.pep
Perfect Score: 1561
Sequence: 1 MASGQGPGPPRQECGEPALP.....LVVLGVVLLGQFVVRRFFKS 211

Scoring table: PAM 150
Gap 11

Searched: 59021 seqs, 21210388 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: swiss-prot34
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11

Statistics: Mean 47.042; Variance 90.697; scale 0.519

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description	Pred. No.
1	274	17.6	190	1	BCLX_CHICK	APOPTOSIS REGULATOR B	3.45e-30
2	270	17.3	233	1	BCLX_RAT	APOPTOSIS REGULATOR B	1.95e-29
3	266	17.0	233	1	BCLX_HUMAN	APOPTOSIS REGULATOR B	1.10e-28
4	261	16.7	233	1	BCL2_CHICK	APOPTOSIS REGULATOR B	9.46e-28
5	255	16.3	205	1	BC2B_HUMAN	PROTEIN BCL-2-BETA.	1.24e-26
6	255	16.3	239	1	BC2A_HUMAN	PROTEIN BCL-2-ALPHA.	1.24e-26
7	250	16.0	199	1	BC2B_MOUSE	PROTEIN BCL-2-BETA.	1.05e-25
8	250	16.0	236	1	BC2A_MOUSE	PROTEIN BCL-2 ALPHA.	1.05e-25
9	248	15.9	236	1	BCL2_RAT	PROTEIN BCL-2 ALPHA.	2.45e-25
10	218	14.0	192	1	BAXA_MOUSE	APOPTOSIS REGULATOR B	7.08e-20
11	206	13.2	192	1	BAXA_HUMAN	APOPTOSIS REGULATOR B	9.65e-18
12	201	12.9	218	1	BAXB_HUMAN	APOPTOSIS REGULATOR B	7.31e-17
13	192	12.3	350	6	MCL1_HUMAN	INDUCED MYELOID LEUKE	2.70e-15
14	177	11.3	143	1	BAXD_HUMAN	BAX PROTEIN, CYTOPLAS	9.88e-13
15	176	11.3	172	5	HSA1_MOUSE	HEMOPOIETIC-SPECIFIC	1.46e-12
16	140	9.0	179	3	EAR_ASFE4	APOPTOSIS REGULATOR B	9.67e-07
17	137	8.8	179	3	EAR_ASFB7	APOPTOSIS REGULATOR B	2.79e-06
18	134	8.6	179	3	EAR_ASFM2	APOPTOSIS REGULATOR B	7.95e-06
19	114	7.3	626	5	HTPG_BACSU	HEAT SHOCK PROTEIN HT	6.32e-03
20	112	7.2	255	2	COX3_CAEEL	CYTOCHROME C OXIDASE	1.19e-02
21	107	6.9	470	9	THRC_PSEAE	THREONINE SYNTHASE (E	5.64e-02
22	106	6.8	133	4	GOP_BPP4	GOP PROTEIN.	7.65e-02
23	100	6.4	567	5	HXT3_YEAST	LOW-AFFINITY GLUCOSE	4.58e-01
24	100	6.4	576	5	HXT4_YEAST	LOW-AFFINITY GLUCOSE	4.58e-01
25	97	6.2	814	4	GUNE_CLOTM	ENDOGLUCANASE E PRECU	1.09e+00
26	95	6.1	206	7	PEM2_YEAST	METHYLENE-FATTY-ACYL-	1.92e+00
27	96	6.1	527	11	YAD7_SCHPO	HYPOTHETICAL 59.6 KD	1.45e+00
28	95	6.1	569	5	HXT8_YEAST	HEXOSE TRANSPORTER HX	1.92e+00
29	95	6.1	817	8	RRPO_TBSVC	PROBABLE RNA-DIRECTED	1.92e+00
30	95	6.1	885	1	AR56_SCHPO	ARG11 PROTEIN PRECURS	1.92e+00
31	94	6.0	531	11	YP60_METTH	HYPOTHETICAL 60.5 KD	2.54e+00
32	94	6.0	534	4	FMO2_RABIT	DIMETHYLANILINE MONOO	2.54e+00
33	93	6.0	833	3	CW41_YEAST	PROTEIN CWH41 PRECURS	3.35e+00
34	90	5.8	196	6	MLEF_HUMAN	MYOSIN LIGHT CHAIN 1,	7.57e+00
35	90	5.8	294	6	MORA_PSEPU	MORPHINE 6-DEHYDROGEN	7.57e+00
36	90	5.8	567	5	HXTA_YEAST	HEXOSE TRANSPORTER HX	7.57e+00
37	90	5.8	567	5	HXT9_YEAST	HEXOSE TRANSPORTER HX	7.57e+00
38	90	5.8	570	5	HXT6_YEAST	HIGH AFFINITY HEXOSE	7.57e+00
39	90	5.8	570	5	HXT7_YEAST	HIGH AFFINITY HEXOSE	7.57e+00
40	90	5.8	680	8	PTFB_MYCGE	PTS SYSTEM, FRUCTOSE-	7.57e+00
41	91	5.8	822	11	YJC9_YEAST	HYPOTHETICAL 95.4 KD	5.78e+00
42	90	5.8	1103	8	RPOB_CYAPA	DNA-DIRECTED RNA POLY	7.57e+00
43	89	5.7	412	11	YAG7_SCHPO	HYPOTHETICAL 45.7 KD	9.89e+00
44	89	5.7	441	1	AROA_PASMU	3-PHOSPHOSHIKIMATE 1-	9.89e+00
45	89	5.7	482	9	TISD_HUMAN	TIS11D PROTEIN (EGF-R	9.89e+00

ALIGNMENTS

RESULT 1
ID BCLX_CHICK STANDARD; PRT; 190 AA.
AC Q07816;
DT 01-FEB-1995 (REL. 31, CREATED)
DT 01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)
DT 01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)
DE APOPTOSIS REGULATOR BCL-X.
GN BCL-X.
OS GALLUS GALLUS (CHICKEN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; AVES; NEOGNATHAE;
OC GALLIFORMES.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 93364977.
RA BOISE L.H., GONZALEZ-GARCIA M., POSTEMA C.E., DING L., LINDSTEN T.,
RA TURKA L.A., MAO X., NUNEZ G., THOMPSON C.B.;
RL CELL 74:597-608(1993).
CC -!- FUNCTION: DOMINANT REGULATOR OF APOPTOTIC CELL DEATH.
CC -!- TISSUE SPECIFICITY: HIGHEST EXPRESSION IN ORGANS WITH
CC LYMPHOID DEVELOPMENT.
CC -!- SIMILARITY: BELONGS TO THE BCL-2 FAMILY.
DR EMBL; Z23110; G510899; -.
DR PIR; A47537; A47537.
DR PROSITE; PS01080; BCL2.
KW APOPTOSIS.
SQ SEQUENCE 190 AA; 21467 MW; 7874E430 CRC32;

Query Match 17.6%; Score 274; DB 1; Length 190;
Best Local Similarity 28.2%; Pred. No. 3.45e-30;
Matches 37; Conservative 33; Mismatches 59; Indels 2; Gaps 2;

```
Db     60 vvngatvhrsslevheivrasdvrqalrdagdefelryrrafsdltsqlhitpgtayqsf 119
          |  |     :| ::     ::| : |   ||::: ||  |  :  :|: |:  ||: |
Qy     52 VAAPADPEMVTLPLQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENAYEYF 111

Db    120 eqvvnelfhdgvnwgrivaffsfggalcvesvdkemrvlvgrivswmttyltdh-ldpwi 178
            :   ||  |:||||:||:::||  | :   :  :  ::|::   :  ::  | :  ||
Qy    112 TKIATSLFESGINWGRVVALLGFGYRLALHVYQHGLTGFLGQVTRFVVDFMLHHCIARWI 171

Db    179 qenggwvrtal 189
           : |||| :||
Qy    172 AQRGGWV-AAL 181
```

RESULT 2
ID BCLX_RAT STANDARD; PRT; 233 AA.

```
AC   P53563;
DT   01-OCT-1996 (REL. 34, CREATED)
DT   01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   APOPTOSIS REGULATOR BCL-X.
GN   BCLX.
OS   RATTUS NORVEGICUS (RAT).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; RODENTIA.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=BRAIN;
RA   MICHAELIDIS T.M.;
RL   SUBMITTED (DEC-1994) TO EMBL/GENBANK/DDBJ DATA BANKS.
CC   -!- FUNCTION: DOMINANT REGULATOR OF APOPTOTIC CELL DEATH.
CC   -!- ALTERNATIVE PRODUCTS: TWO ISOFORMS, BCX-X(L) (SHOWN HERE) AND
CC       BCL-X(S), ARE DERIVED BY ALTERNATIVE SPLICING.
CC   -!- SIMILARITY: BELONGS TO THE BCL-2 FAMILY.
DR   EMBL; X82537; G607177; -.
DR   EMBL; X82537; G607178; -.
KW   APOPTOSIS; ALTERNATIVE SPLICING.
FT   VARSPLIC    126    188       MISSING (IN BCL-X(S)).
SQ   SEQUENCE   233 AA;  26130 MW;  E0589815 CRC32;

  Query Match           17.3%;  Score 270;  DB 1;  Length 233;
  Best Local Similarity  25.5%;  Pred. No. 1.95e-29;
  Matches    36;  Conservative   42;  Mismatches  60;  Indels   3;  Gaps   3;

Db     61 dspavngatghs-ssldarevipmaavkqalreagdefelryrrafsdltsqlhitpgta 119
          :: :| :::     :|  :    |: | : |   ||::: ||   |  :  :|: |:  |
Qy     48 EAEGVAAPADPEMVTLPLQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENA 107

Db    120 yqsfeqvvnelfrdgvnwgrivaffsfggalcvesvdkemqvlvsriaswmatylndh-l 178
          |: |  :   ||  |:||||:||:::||  | :   :  :  ::::::  :  ::  | :
Qy    108 YEYFTKIATSLFESGINWGRVVALLGFGYRLALHVYQHGLTGFLGQVTRFVVDFMLHHCI 167

Db    179 epwiqenggwdtfvdlygnna 199
            || : ||| : ::| || :
Qy    168 ARWIAQRGGWVAALNL-GNGP 187


RESULT    3
ID   BCLX_HUMAN     STANDARD;      PRT;   233 AA.
AC   Q07817;
DT   01-FEB-1995 (REL. 31, CREATED)
DT   01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   APOPTOSIS REGULATOR BCL-X.
GN   BCLX.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 93364977.
RA   BOISE L.H., GONZALEZ-GARCIA M., POSTEMA C.E., DING L.,
RA   LINDSTEN T., TURKA L.A., MAO X., NUNEZ G., THOMPSON C.B.;
RL   CELL 74:597-608(1993).
CC   -!- FUNCTION: DOMINANT REGULATOR OF APOPTOTIC CELL DEATH.
CC   -!- ALTERNATIVE PRODUCTS: TWO ISOFORMS, BCX-X(L) (SHOWN HERE) AND
CC       BCL-X(S), ARE DERIVED BY ALTERNATIVE SPLICING.
CC   -!- TISSUE SPECIFICITY: BCL-X(S) IS EXPRESSED AT HIGH LEVELS IN CELLS
CC       THAT UNDERGO A HIGH RATE OF TURNOVER, SUCH AS DEVELOPING
CC       LYMPHOCYTES. IN CONTRAST, BCL-X(L) IS FOUND IN TISSUES CONTAINING
CC       LONG-LIVED POSTMITOTIC CELLS, SUCH AS ADULT BRAIN.
CC   -!- SIMILARITY: BELONGS TO THE BCL-2 FAMILY.
DR   EMBL; Z23116; G623237; -.
DR   EMBL; Z23115; G510901; -.
DR   MIM; 600039; -.
DR   PROSITE; PS01080; BCL2.
KW   APOPTOSIS; ALTERNATIVE SPLICING.
FT   VARSPLIC    126    188       MISSING (IN BCL-X(S)).
FT   CONFLICT     70     70       G -> A (IN G510901).
SQ   SEQUENCE   233 AA;  26049 MW;  57C67491 CRC32;

  Query Match           17.0%;  Score 266;  DB 1;  Length 233;
  Best Local Similarity  25.5%;  Pred. No. 1.10e-28;
  Matches    36;  Conservative   42;  Mismatches  60;  Indels   3;  Gaps   3;

Db     61 dspavngatghs-ssldarevipmaavkqalreagdefelryrrafsdltsqlhitpgta 119
          :: :| :::     :|  :    |: | : |   ||::: ||   |  :  :|: |:  |
Qy     48 EAEGVAAPADPEMVTLPLQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENA 107

Db    120 yqsfeqvvnelfrdgvnwgrivaffsfggalcvesvdkemqvlvsriaawmatylndh-l 178
          |: |  :   ||  |:||||:||:::||  | :   :  :  ::::::  :  ::  | :
Qy    108 YEYFTKIATSLFESGINWGRVVALLGFGYRLALHVYQHGLTGFLGQVTRFVVDFMLHHCI 167

Db    179 epwiqenggwdtfvelygnna 199
            || : ||| : ::| || :
Qy    168 ARWIAQRGGWVAALNL-GNGP 187


RESULT    4
ID   BCL2_CHICK     STANDARD;      PRT;   233 AA.
AC   Q00709;
DT   01-APR-1993 (REL. 25, CREATED)
DT   01-APR-1993 (REL. 25, LAST SEQUENCE UPDATE)
DT   01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE   APOPTOSIS REGULATOR BCL-2.
GN   BCL-2.
OS   GALLUS GALLUS (CHICKEN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; AVES; NEOGNATHAE;
OC   GALLIFORMES.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 92375724.
RA   EGUCHI Y., EWERT D.L., TSUJIMOTO Y.;
RL   NUCLEIC ACIDS RES. 20:4187-4192(1992).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 92379084.
RA   CAZALS-HATEM D.L., LOUIE D.C., TANAKA S., REED J.C.;
RL   BIOCHIM. BIOPHYS. ACTA 1132:109-113(1992).
CC   -!- FUNCTION: PROLONGS THE SURVIVAL OF HEMATOPOIETIC CELLS IN THE
CC       ABSENCE OF REQUIRED GROWTH FACTORS AND ALSO IN THE PRESENCE OF
CC       VARIOUS STIMULI INDUCING CELLULAR DEATH. BLOCKS APOPTOSIS. MIGHT
CC       FUNCTION IN AN ANTIOXIDANT PATHWAY TO PREVENT APOPTOSIS AT SITES
CC       OF FREE RADICAL GENERATION SUCH AS MITOCHONDRIA.
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL INNER MEMBRANE
CC       (BY SIMILARITY).
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN A VARIETY OF TISSUES.
CC   -!- SIMILARITY: BELONGS TO THE BCL-2 FAMILY.
DR   EMBL; D11382; G222794; -.
DR   EMBL; D11381; G222794; JOINED.
DR   EMBL; Z11961; G62970; -.
DR   PIR; A37332; A37332.
DR   PIR; S24390; S24390.
DR   PROSITE; PS01080; BCL2.
KW   APOPTOSIS; TRANSMEMBRANE; MITOCHONDRION.
FT   TRANSMEM    208    228       POTENTIAL.
FT   CONFLICT     64     64       E -> S (IN REF. 2).
FT   CONFLICT     67     82       GSAAASEVPPAEGLRP -> ARLLLVRCPRLRGCA
FT                                (IN REF. 2).
FT   CONFLICT    121    121       H -> T (IN REF. 2).
FT   CONFLICT    139    139       G -> V (IN REF. 2).
SQ   SEQUENCE   233 AA;  25687 MW;  3376502C CRC32;

  Query Match           16.7%;  Score 261;  DB 1;  Length 233;
  Best Local Similarity  28.5%;  Pred. No. 9.46e-28;
  Matches    41;  Conservative   41;  Mismatches  57;  Indels   5;  Gaps   5;

Db     60 hhrpeppgsaaasevppae-glrpapp-g-vhlalrqagdefsrryqrdfaqmsgqlhlt 116
          ::  |: | ||:::       |:|::  | |   |   ||:::|||: :|  |  :|: |
Qy     44 QQEQEAEGVAAPADPEMVTLPLQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPT 103

Db    117 pftahgrfvavveelfrdgvnwgrivaffefggvmcvesvnremsplvdniatwmteyln 176
```

```
          :  |   |  :   ||  |:||||:||:: ||  : :    : :: ::  ::  : :::
Qy    104 AENAYEYFTKIATSLFESGINWGRVVALLGFGYRLALHVYQHGLTGFLGQVTRFVVDFML 163

Db    177 rh-lhnwiqdnggwdafvelygns 199
          :| :  || : ||| | ::| ||:
Qy    164 HHCIARWIAQRGGWVAALNL-GNG 186


RESULT    5
ID   BC2B_HUMAN     STANDARD;      PRT;   205 AA.
AC   P10416;
DT   01-MAR-1989 (REL. 10, CREATED)
DT   01-APR-1993 (REL. 25, LAST SEQUENCE UPDATE)
DT   01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE   PROTEIN BCL-2-BETA.
GN   BCL2.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 86259760.
RA   TSUJIMOTO Y., CROCE C.M.;
RL   PROC. NATL. ACAD. SCI. U.S.A. 83:5214-5218(1986).
RN   [2]
RP   REVISIONS TO 96 AND 110.
RX   MEDLINE; 92375724.
RA   EGUCHI Y., EWERT D.L., TSUJIMOTO Y.;
RL   NUCLEIC ACIDS RES. 20:4187-4192(1992).
RN   [3]
RP   SUBCELLULAR LOCATION.
RX   MEDLINE; 91066924.
RA   HOCKENBERY D., NUNEZ G., MILLIMAN C., SCHREIBER R.D., KORSMEYER S.J.;
RL   NATURE 348:334-336(1990).
CC   -!- FUNCTION: PROLONGS THE SURVIVAL OF HEMATOPOIETIC CELLS IN THE
CC       ABSENCE OF REQUIRED GROWTH FACTORS AND ALSO IN THE PRESENCE OF
CC       VARIOUS STIMULI INDUCING CELLULAR DEATH. BLOCKS APOPTOSIS. MIGHT
CC       FUNCTION IN AN ANTIOXIDANT PATHWAY TO PREVENT APOPTOSIS AT SITES
CC       OF FREE RADICAL GENERATION SUCH AS MITOCHONDRIA.
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL INNER MEMBRANE.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN A VARIETY OF TISSUES.
CC   -!- DISEASE: INVOLVED IN FOLLICULAR LYMPHOMA (FL) (ALSO KNOWN AS TYPE
CC       II CHRONIC LYMPHATIC LEUKEMIA) BY A CHROMOSOMAL TRANSLOCATION
CC       T(14;18)(Q32;Q21) WHICH INVOLVES BCL2 AND IMMUNOGLOBULIN GENE
CC       REGIONS.
CC   -!- ALTERNATIVE PRODUCTS: TWO FORMS OF BCL-2: ALPHA, AND BETA, ARE
CC       PRODUCED BY ALTERNATIVE SPLICING OF THE SAME GENE. THEY ONLY
CC       DIFFER AT THEIR C-TERMINAL ENDS.
CC   -!- SIMILARITY: BELONGS TO THE BCL-2 FAMILY.
DR   EMBL; M13995; G179369; ALT_SEQ.
DR   PIR; B29409; TVHUB1.
DR   PIR; D37332; D37332.
DR   MIM; 151430; -.
DR   PROSITE; PS01080; BCL2.
KW   PROTO-ONCOGENE; APOPTOSIS; ALTERNATIVE SPLICING; MEMBRANE;
KW   MITOCHONDRION; CHROMOSOMAL TRANSLOCATION.
SQ   SEQUENCE   205 AA;  22311 MW;  ED321E5E CRC32;

  Query Match           16.3%;  Score 255;  DB 1;  Length 205;
  Best Local Similarity  29.7%;  Pred. No. 1.24e-26;
  Matches    38;  Conservative   35;  Mismatches  51;  Indels   4;  Gaps   3;

Db     74 tpaapgaaagpalspvppvvhlt--lrqagddfsrryrrdfaemssqlhltpftargrfa 131
          :|| |   : | | | : : ::   |   |||::||| :|  |  :|: |:  |   |:
Qy     54 APADPEMVTLP-LQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENAYEYFT 112

Db    132 tvveelfrdgvnwgrivaffefggvmcvesvnremsplvdnialwmteylnrh-lhtwiq 190
          :   ||  |:||||:||:: ||  : :    : :: ::  ::  : ::: :| :  ||
Qy    113 KIATSLFESGINWGRVVALLGFGYRLALHVYQHGLTGFLGQVTRFVVDFMLHHCIARWIA 172

Db    191 dnggwvga 198
          : ||||:|
Qy    173 QRGGWVAA 180
```

```
RESULT    6
ID   BC2A_HUMAN     STANDARD;      PRT;   239 AA.
AC   P10415;
DT   01-MAR-1989 (REL. 10, CREATED)
DT   01-APR-1993 (REL. 25, LAST SEQUENCE UPDATE)
DT   01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE   PROTEIN BCL-2-ALPHA.
GN   BCL2.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 86259760.
RA   TSUJIMOTO Y., CROCE C.M.;
RL   PROC. NATL. ACAD. SCI. U.S.A. 83:5214-5218(1986).
RN   [2]
RP   REVISIONS TO 96; 110 AND 237.
RX   MEDLINE; 92375724.
RA   EGUCHI Y., EWERT D.L., TSUJIMOTO Y.;
RL   NUCLEIC ACIDS RES. 20:4187-4192(1992).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 87002488.
RA   CLEARY M.L., SMITH S.D., SKLAR J.;
RL   CELL 47:19-28(1986).
RN   [4]
RP   SUBCELLULAR LOCATION.
RX   MEDLINE; 91066924.
RA   HOCKENBERY D., NUNEZ G., MILLIMAN C., SCHREIBER R.D., KORSMEYER S.J.;
RL   NATURE 348:334-336(1990).
CC   -!- FUNCTION: PROLONGS THE SURVIVAL OF HEMATOPOIETIC CELLS IN THE
CC       ABSENCE OF REQUIRED GROWTH FACTORS AND ALSO IN THE PRESENCE OF
CC       VARIOUS STIMULI INDUCING CELLULAR DEATH. BLOCKS APOPTOSIS. MIGHT
CC       FUNCTION IN AN ANTIOXIDANT PATHWAY TO PREVENT APOPTOSIS AT SITES
CC       OF FREE RADICAL GENERATION SUCH AS MITOCHONDRIA.
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL INNER MEMBRANE.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN A VARIETY OF TISSUES.
CC   -!- DISEASE: INVOLVED IN FOLLICULAR LYMPHOMA (FL) (ALSO KNOWN AS TYPE
CC       II CHRONIC LYMPHATIC LEUKEMIA) BY A CHROMOSOMAL TRANSLOCATION
CC       T(14;18)(Q32;Q21) WHICH INVOLVES BCL2 AND IMMUNOGLOBULIN GENE
CC       REGIONS.
CC   -!- ALTERNATIVE PRODUCTS: TWO FORMS OF BCL-2: ALPHA, AND BETA, ARE
CC       PRODUCED BY ALTERNATIVE SPLICING OF THE SAME GENE. THEY ONLY
CC       DIFFER AT THEIR C-TERMINAL ENDS.
CC   -!- SIMILARITY: BELONGS TO THE BCL-2 FAMILY.
DR   EMBL; M13994; G179367; ALT_SEQ.
DR   EMBL; M14745; G179371; -.
DR   PIR; A29409; TVHUA1.
DR   PIR; A24428; TVHUBC.
DR   PIR; C37332; C37332.
DR   MIM; 151430; -.
DR   PROSITE; PS01080; BCL2.
KW   PROTO-ONCOGENE; APOPTOSIS; ALTERNATIVE SPLICING; TRANSMEMBRANE;
KW   MITOCHONDRION; CHROMOSOMAL TRANSLOCATION.
FT   TRANSMEM    212    233       POTENTIAL.
FT   CONFLICT     59     59       P -> T (IN REF. 3).
FT   CONFLICT    117    117       S -> R (IN REF. 3).
SQ   SEQUENCE   239 AA;  26266 MW;  75084B59 CRC32;

  Query Match           16.3%;  Score 255;  DB 1;  Length 239;
  Best Local Similarity  29.3%;  Pred. No. 1.24e-26;
  Matches    41;  Conservative   37;  Mismatches  57;  Indels   5;  Gaps   4;

Db     74 tpaapgaaagpalspvppvvhlt--lrqagddfsrryrrdfaemssqlhltpftargrfa 131
          :|| |   : | | | : : ::   |   |||::||| :|  |  :|: |:  |   |:
Qy     54 APADPEMVTLP-LQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENAYEYFT 112

Db    132 tvveelfrdgvnwgrivaffefggvmcvesvnremsplvdnialwmteylnrh-lhtwiq 190
          :   ||  |:||||:||:: ||  : :    : :: ::  ::  : ::: :| :  ||
Qy    113 KIATSLFESGINWGRVVALLGFGYRLALHVYQHGLTGFLGQVTRFVVDFMLHHCIARWIA 172
```

```
Db    191 dnggwdafvel-ygpsmrpl 209
          : ||| | ::|  || :  |
Qy    173 QRGGWVAALNLGNGPILNVL 192


RESULT    7
ID    BC2B_MOUSE     STANDARD;      PRT;   199 AA.
AC    P10418;
DT    01-MAR-1989 (REL. 10, CREATED)
DT    01-MAR-1989 (REL. 10, LAST SEQUENCE UPDATE)
DT    01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE    PROTEIN BCL-2-BETA.
GN    BCL2 OR BCL-2.
OS    MUS MUSCULUS (MOUSE).
OC    EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC    EUTHERIA; RODENTIA.
RN    [1]
RP    SEQUENCE FROM N.A.
RX    MEDLINE; 87187643.
RA    NEGRINI M., SILINI E., KOZAK C., TSUJIMOTO Y., CROCE C.M.;
RL    CELL 49:455-463(1987).
CC    -!- FUNCTION: PROLONGS THE SURVIVAL OF HEMATOPOIETIC CELLS IN THE
CC        ABSENCE OF REQUIRED GROWTH FACTORS AND ALSO IN THE PRESENCE OF
CC        VARIOUS STIMULI INDUCING CELLULAR DEATH. BLOCKS APOPTOSIS. MIGHT
CC        FUNCTION IN AN ANTIOXIDANT PATHWAY TO PREVENT APOPTOSIS AT SITES
CC        OF FREE RADICAL GENERATION SUCH AS MITOCHONDRIA.
CC    -!- SUBCELLULAR LOCATION: MITOCHONDRIAL INNER MEMBRANE.
CC    -!- TISSUE SPECIFICITY: EXPRESSED IN A VARIETY OF TISSUES.
CC    -!- ALTERNATIVE PRODUCTS: TWO FORMS OF BCL-2: ALPHA, AND BETA, ARE
CC        PRODUCED BY ALTERNATIVE SPLICING OF THE SAME GENE. THEY ONLY
CC        DIFFER AT THEIR C-TERMINAL ENDS.
CC    -!- SIMILARITY: BELONGS TO THE BCL-2 FAMILY.
DR    EMBL; M16506; G387110; -.
DR    PIR; B25960; TVMSB1.
DR    PROSITE; PS01080; BCL2.
KW    APOPTOSIS; ALTERNATIVE SPLICING; MEMBRANE; MITOCHONDRION.
SQ    SEQUENCE   199 AA;  22299 MW;  3E60FA47 CRC32;

  Query Match           16.0%;  Score 250;  DB 1;  Length 199;
  Best Local Similarity  31.1%;  Pred. No. 1.05e-25;
  Matches    32;  Conservative   29;  Mismatches  41;  Indels   1;  Gaps   1;

Db     94 lrragddfsrryrrdfaemssqlhltpftargrfatvveelfrdgvnwgrivaffefggv 153
          |   |||::|||  :|  |  :|: |:  |   |: :   ||  |:||||:||:: ||
Qy     78 LAIIGDDINRRYDSEFQTMLQHLQPTAENAYEYFTKIATSLFESGINWGRVVALLGFGYR 137

Db    154 mcvesvnremsplvdnialwmteylnrh-lhtwiqdnggwvga 195
          : :    : :: ::  ::  : ::: :| :  || : ||||:|
Qy    138 LALHVYQHGLTGFLGQVTRFVVDFMLHHCIARWIAQRGGWVAA 180


RESULT    8
ID    BC2A_MOUSE     STANDARD;      PRT;   236 AA.
AC    P10417;
DT    01-MAR-1989 (REL. 10, CREATED)
DT    01-APR-1993 (REL. 25, LAST SEQUENCE UPDATE)
DT    01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE    PROTEIN BCL-2 ALPHA.
GN    BCL2 OR BCL-2.
OS    MUS MUSCULUS (MOUSE).
OC    EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC    EUTHERIA; RODENTIA.
RN    [1]
RP    SEQUENCE FROM N.A.
RC    STRAIN=BALB/C; TISSUE=LIVER;
RX    MEDLINE; 87187643.
RA    NEGRINI M., SILINI E., KOZAK C., TSUJIMOTO Y., CROCE C.M.;
RL    CELL 49:455-463(1987).
RN    [2]
RP    REVISIONS TO 221-222.
RX    MEDLINE; 92375724.
RA    EGUCHI Y., EWERT D.L., TSUJIMOTO Y.;
RL    NUCLEIC ACIDS RES. 20:4187-4192(1992).
CC    -!- FUNCTION: PROLONGS THE SURVIVAL OF HEMATOPOIETIC CELLS IN THE
CC        ABSENCE OF REQUIRED GROWTH FACTORS AND ALSO IN THE PRESENCE OF
CC        VARIOUS STIMULI INDUCING CELLULAR DEATH. BLOCKS APOPTOSIS. MIGHT
CC        FUNCTION IN AN ANTIOXIDANT PATHWAY TO PREVENT APOPTOSIS AT SITES
CC        OF FREE RADICAL GENERATION SUCH AS MITOCHONDRIA.
CC    -!- SUBCELLULAR LOCATION: MITOCHONDRIAL INNER MEMBRANE.
CC    -!- TISSUE SPECIFICITY: EXPRESSED IN A VARIETY OF TISSUES.
CC    -!- ALTERNATIVE PRODUCTS: TWO FORMS OF BCL-2: ALPHA, AND BETA, ARE
CC        PRODUCED BY ALTERNATIVE SPLICING OF THE SAME GENE. THEY ONLY
CC        DIFFER AT THEIR C-TERMINAL ENDS.
CC    -!- SIMILARITY: BELONGS TO THE BCL-2 FAMILY.
DR    EMBL; L31532; G387109; -.
DR    EMBL; M16506; G387109; JOINED.
DR    PIR; A25960; TVMSA1.
DR    PIR; E37332; E37332.
DR    PROSITE; PS01080; BCL2.
KW    APOPTOSIS; ALTERNATIVE SPLICING; TRANSMEMBRANE; MITOCHONDRION.
FT    TRANSMEM    209    230       POTENTIAL.
SQ    SEQUENCE   236 AA;  26425 MW;  7ADFE975 CRC32;

  Query Match           16.0%;  Score 250;  DB 1;  Length 236;
  Best Local Similarity  30.4%;  Pred. No. 1.05e-25;
  Matches    35;  Conservative   31;  Mismatches  47;  Indels   2;  Gaps   2;

Db     94 lrragddfsrryrrdfaemssqlhltpftargrfatvveelfrdgvnwgrivaffefggv 153
          |   |||::|||  :|  |  :|: |:  |   |: :   ||  |:||||:||:: ||
Qy     78 LAIIGDDINRRYDSEFQTMLQHLQPTAENAYEYFTKIATSLFESGINWGRVVALLGFGYR 137

Db    154 mcvesvnremsplvdnialwmteylnrh-lhtwiqdnggwdafvel-ygpsmrpl 206
          : :    : :: ::  ::  : ::: :| :  || : ||| | ::|  || :  |
Qy    138 LALHVYQHGLTGFLGQVTRFVVDFMLHHCIARWIAQRGGWVAALNLGNGPILNVL 192


RESULT    9
ID    BCL2_RAT       STANDARD;      PRT;   236 AA.
AC    P49950;
DT    01-OCT-1996 (REL. 34, CREATED)
DT    01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT    01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE    PROTEIN BCL-2 ALPHA.
GN    BCL2 OR BCL-2.
OS    RATTUS NORVEGICUS (RAT).
OC    EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC    EUTHERIA; RODENTIA.
RN    [1]
RP    SEQUENCE FROM N.A.
RC    TISSUE=BRAIN;
RX    MEDLINE; 94193015.
RA    SATO T., IRIE S., KRAJEWSKI S., REED J.C.;
RL    GENE 140:291-292(1994).
CC    -!- FUNCTION: PROLONGS THE SURVIVAL OF HEMATOPOIETIC CELLS IN THE
CC        ABSENCE OF REQUIRED GROWTH FACTORS AND ALSO IN THE PRESENCE OF
CC        VARIOUS STIMULI INDUCING CELLULAR DEATH. BLOCKS APOPTOSIS. MIGHT
CC        FUNCTION IN AN ANTIOXIDANT PATHWAY TO PREVENT APOPTOSIS AT SITES
CC        OF FREE RADICAL GENERATION SUCH AS MITOCHONDRIA.
CC    -!- SUBCELLULAR LOCATION: MITOCHONDRIAL INNER MEMBRANE.
CC    -!- TISSUE SPECIFICITY: EXPRESSED IN A VARIETY OF TISSUES (BY
CC        SIMILARITY).
CC    -!- ALTERNATIVE PRODUCTS: TWO FORMS OF BCL-2: ALPHA, AND BETA, ARE
CC        PRODUCED BY ALTERNATIVE SPLICING OF THE SAME GENE. THEY ONLY
CC        DIFFER AT THEIR C-TERMINAL ENDS.
CC    -!- SIMILARITY: BELONGS TO THE BCL-2 FAMILY.
DR    EMBL; L14680; G408947; -.
KW    APOPTOSIS; ALTERNATIVE SPLICING; TRANSMEMBRANE; MITOCHONDRION.
FT    TRANSMEM    209    230       POTENTIAL.
SQ    SEQUENCE   236 AA;  26550 MW;  336E6B40 CRC32;

  Query Match           15.9%;  Score 248;  DB 1;  Length 236;
  Best Local Similarity  27.2%;  Pred. No. 2.45e-25;
  Matches    40;  Conservative   43;  Mismatches  60;  Indels   4;  Gaps   4;

Db     63 rdtaartsplrp-lvanagpalspvppvvhltlrragddfsrryrrdfaemssqlhltpf 121
```

```
          ::: : ::|  | :|: :  : | :  | :  |  |||::|||  :|  |  :|: |:
Qy     47 QEAEGVAAPADPEMVTLPLQPSSTMGQVGR-QLAIIGDDINRRYDSEFQTMLQHLQPTAE 105

Db    122 targrfatvveelfrdgvnwgrivaffefggvmcvgsvnremsplvdnialwmteylnrh 181
          |   |: :   ||  |:||||:||:: ||  : :    : :: ::  ::  : ::: :|
Qy    106 NAYEYFTKIATSLFESGINWGRVVALLGFGYRLALHVYQHGLTGFLGQVTRFVVDFMLHH 165

Db    182 -lhtwiqdnggwdafvel-ygpsmrpl 206
           :  || : ||| | ::|  || :  |
Qy    166 CIARWIAQRGGWVAALNLGNGPILNVL 192


RESULT   10
ID   BAXA_MOUSE     STANDARD;      PRT;   192 AA.
AC   Q07813;
DT   01-FEB-1995 (REL. 31, CREATED)
DT   01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   APOPTOSIS REGULATOR BAX, MEMBRANE ISOFORM ALPHA.
GN   BAX.
OS   MUS MUSCULUS (MOUSE).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; RODENTIA.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6 X DBA/2 F1;
RX   MEDLINE; 93364978.
RA   OLTVAI Z.N., MILLIMAN C.L., KORSMEYER S.J.;
RL   CELL 74:609-619(1993).
CC   -!- FUNCTION: ACCELERATES PROGRAMMED CELL DEATH BY BINDING TO THE
CC       APOPTOSIS REPRESSOR BCL-2.
CC   -!- ALTERNATIVE PRODUCTS: A 21 KD MEMBRANE PROTEIN ALPHA AND THE TWO
CC       CYTOPLASMIC PROTEINS BETA AND GAMMA ARE GENERATED BY ALTERNATIVE
CC       SPLICING.
CC   -!- SUBUNIT: FORMS HOMODIMERS AND HETERODIMERS TOGETHER WITH BCL-2.
CC   -!- SUBCELLULAR LOCATION: MEMBRANE.
CC   -!- SIMILARITY: BELONGS TO THE BCL-2 FAMILY.
DR   EMBL; L22472; G388192; -.
DR   PROSITE; PS01080; BCL2.
KW   APOPTOSIS; TRANSMEMBRANE; ALTERNATIVE SPLICING.
FT   TRANSMEM    172    192       POTENTIAL.
SQ   SEQUENCE   192 AA;  21394 MW;  BD035304 CRC32;

  Query Match           14.0%;  Score 218;  DB 1;  Length 192;
  Best Local Similarity  25.6%;  Pred. No. 7.08e-20;
  Matches    46;  Conservative   50;  Mismatches  75;  Indels   9;  Gaps   8;

Db     16 seqimktgafllqgfiqdragrmagetpeltleqppqdastk-klseclrrigdelds-- 72
          :|:: ::  |  :   |:  |  |   ||: :  | | :||:  ::  |  |||:::
Qy     31 TEEVFRSYVFYRHQQEQEAEGVAAPADPEM-VTLPLQPSSTMGQVGRQLAIIGDDINRRY 89

Db     73 nmelqrmiadvdtdspre-vff-rvaadmfadgnfnwgrvvalfyfasklvlkalctkvp 130
          : |:| |:  ::  :     :| ::|: :|  | :||||||||: |: :| |       :
Qy     90 DSEFQTMLQHLQPTAENAYEYFTKIATSLFESG-INWGRVVALLGFGYRLALHVYQHGLT 148

Db    131 elirtimgwtldf-lrerllvwiqdqggwegllsyfgtptwqtvtifvagvltasltiwk 189
          ::  :    :|| |:  :  || ::||| : |:    |    | : :: || : : :::
Qy    149 GFLGQVTRFVVDFMLHHCIARWIAQRGGWVAALNLGNGPI-LNVLVVLGVVLLGQFVVRR 207


RESULT   11
ID   BAXA_HUMAN     STANDARD;      PRT;   192 AA.
AC   Q07812;
DT   01-FEB-1995 (REL. 31, CREATED)
DT   01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   APOPTOSIS REGULATOR BAX, MEMBRANE ISOFORM ALPHA.
GN   BAX.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=B-CELL;
RX   MEDLINE; 93364978.
RA   OLTVAI Z.N., MILLIMAN C.L., KORSMEYER S.J.;
RL   CELL 74:609-619(1993).
CC   -!- FUNCTION: ACCELERATES PROGRAMMED CELL DEATH BY BINDING TO THE
CC       APOPTOSIS REPRESSOR BCL-2.
CC   -!- SUBUNIT: FORMS HOMODIMERS AND HETERODIMERS TOGETHER WITH BCL-2.
CC   -!- SUBCELLULAR LOCATION: MEMBRANE.
CC   -!- ALTERNATIVE PRODUCTS: THE MEMBRANE ISOFORM ALPHA AND THE THREE
CC       CYTOPLASMIC ISOFORMS, BETA, GAMMA AND DELTA ARE GENERATED BY
CC       ALTERNATIVE SPLICING.
CC   -!- SIMILARITY: BELONGS TO THE BCL-2 FAMILY.
DR   EMBL; L22473; G388166; -.
DR   PIR; A47538; A47538.
DR   MIM; 600040; -.
DR   PROSITE; PS01080; BCL2.
KW   APOPTOSIS; TRANSMEMBRANE; ALTERNATIVE SPLICING.
FT   TRANSMEM    172    192       POTENTIAL.
SQ   SEQUENCE   192 AA;  21184 MW;  B2E6148A CRC32;

  Query Match           13.2%;  Score 206;  DB 1;  Length 192;
  Best Local Similarity  24.5%;  Pred. No. 9.65e-18;
  Matches    40;  Conservative   45;  Mismatches  71;  Indels   7;  Gaps   6;

Db     32 qdragrmggeapelaldpvpqdastkklseclkrigdelds--nmelqrmiaavdtdspr 89
          |:  |  :   ||:   |:   ::   ::  |  |||:::   : |:| |:  ::  :
Qy     47 QEAEGVAAPADPEMVTLPLQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAEN 106

Db     90 e-vff-rvaadmfsdgnfnwgrvvalfyfasklvlkalctkvpelirtimgwtldf-lre 146
             :| ::|: :|  | :|||||||: |: :| |      :  ::  :    :|| |:
Qy    107 AYEYFTKIATSLFESG-INWGRVVALLGFGYRLALHVYQHGLTGFLGQVTRFVVDFMLHH 165

Db    147 rllgwiqdqggwdgllsyfgtptwqtvtifvagvltasltiwk 189
           :  || ::||| : |:    |    | : :: || : : :::
Qy    166 CIARWIAQRGGWVAALNLGNGPI-LNVLVVLGVVLLGQFVVRR 207


RESULT   12
ID   BAXB_HUMAN     STANDARD;      PRT;   218 AA.
AC   Q07814;
DT   01-FEB-1995 (REL. 31, CREATED)
DT   01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   APOPTOSIS REGULATOR BAX, CYTOPLASMIC ISOFORM BETA.
GN   BAX.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=B-CELL;
RX   MEDLINE; 93364978.
RA   OLTVAI Z.N., MILLIMAN C.L., KORSMEYER S.J.;
RL   CELL 74:609-619(1993).
CC   -!- FUNCTION: ACCELERATES PROGRAMMED CELL DEATH BY BINDING TO THE
CC       APOPTOSIS REPRESSOR BCL-2.
CC   -!- SUBUNIT: FORMS HOMODIMERS AND HETERODIMERS TOGETHER WITH BCL-2.
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC   -!- ALTERNATIVE PRODUCTS: THE MEMBRANE ISOFORM ALPHA AND THE THREE
CC       CYTOPLASMIC ISOFORMS, BETA, GAMMA AND DELTA ARE GENERATED BY
CC       ALTERNATIVE SPLICING.
CC   -!- SIMILARITY: BELONGS TO THE BCL-2 FAMILY.
DR   EMBL; L22474; G388168; -.
DR   PIR; B47538; B47538.
DR   MIM; 600040; -.
DR   PROSITE; PS01080; BCL2.
KW   APOPTOSIS; ALTERNATIVE SPLICING.
SQ   SEQUENCE   218 AA;  24220 MW;  82B2FF09 CRC32;

  Query Match           12.9%;  Score 201;  DB 1;  Length 218;
  Best Local Similarity  27.0%;  Pred. No. 7.31e-17;
  Matches    37;  Conservative   36;  Mismatches  58;  Indels   6;  Gaps   5;
```

```
Db     32 qdragrmggeapelaldpvpqdastkklseclkrigdelds--nmelqrmiaavdtdspr 89
          |:  |  :   ||:   |:   ::   ::  |  |||:::   : |:| |:  ::  :
Qy     47 QEAEGVAAPADPEMVTLPLQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAEN 106

Db     90 e-vff-rvaadmfsdgnfnwgrvvalfyfasklvlkalctkvpelirtimgwtldf-lre 146
             :| ::|: :|  | :|||||||: |: :| |      :  ::  :    :|| |:
Qy    107 AYEYFTKIATSLFESG-INWGRVVALLGFGYRLALHVYQHGLTGFLGQVTRFVVDFMLHH 165

Db    147 rllgwiqdqggwvrllk 163
            :  || ::|||| |:
Qy    166 CIARWIAQRGGWVAALN 182


RESULT   13
ID   MCL1_HUMAN     STANDARD;      PRT;   350 AA.
AC   Q07820;
DT   01-FEB-1995 (REL. 31, CREATED)
DT   01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)
DT   01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)
DE   INDUCED MYELOID LEUKEMIA CELL DIFFERENTIATION PROTEIN MCL1.
GN   MCL1.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=MYELOID LEUKEMIA CELLS;
RX   MEDLINE; 93234528.
RA   KOZOPAS K.M., YANG T., BUCHAN H.L., ZHOU P., CRAIG R.W.;
RL   PROC. NATL. ACAD. SCI. U.S.A. 90:3516-3520(1993).
CC   -!- FUNCTION: INVOLVED IN PROGRAMMING OF DIFFERENTIATION AND
CC       CONCOMITANT MAINTENANCE OF VIABILITY BUT NOT OF PROLIFERATION
CC       (PROBABLE).
CC   -!- INDUCTION: EXPRESSION INCREASES EARLY DURING PHORBOL-ESTER INDUCED
CC       DIFFERENTIATION ALONG THE MONOCYTE/MAGROPHAGE PATHWAY IN MYELOID
CC       LEUKEMIA CELL LINES ML-1.
CC   -!- SIMILARITY: BELONGS TO THE BCL2 FAMILY.
DR   EMBL; L08246; -; NOT_ANNOTATED_CDS.
DR   PIR; A47476; A47476.
DR   MIM; 159552; -.
DR   PROSITE; PS01080; BCL2.
KW   APOPTOSIS; TRANSMEMBRANE; DIFFERENTIATION.
FT   UNSURE      227    227       OR A.
FT   TRANSMEM    330    349       POTENTIAL.
SQ   SEQUENCE   350 AA;  37365 MW;  10194B64 CRC32;

  Query Match           12.3%;  Score 192;  DB 6;  Length 350;
  Best Local Similarity  24.3%;  Pred. No. 2.70e-15;
  Matches   34;  Conservative   40;  Mismatches  63;  Indels   3;  Gaps   3;

Db    174 lyrqsleiisrylreqatgakdtkpmgrsgatsrkaletlrrvgdgvqrnhetvfqgmlr 233
          :||:  |  :  :   |  : | |:  |:: :: : : |  :|| : |  :: || ||:
Qy     40 FYRHQQEQEAEGVAAPADPEMVTLPLQPSSTMGQVGRQ-LAIIGDDINRRYDSEFQTMLQ 98

Db    234 kldikneddvkslsrvmihvfsdgvtnwgrivtlisfgafvakhlktinqescieplaes 293
           |:   |:    ::::   :|  |: ||||:|:|::|| :| |:      : :  ::
Qy     99 HLQPTAENAYEYFTKIATSLFESGI-NWGRVVALLGFGYRLALHVYQHGLTGFLGQVTRF 157

Db    294 itd-vlvrtkrdwlvkqrgw 312
          : | :| :    |:  : ||
Qy    158 VVDFMLHHCIARWIAQRGGW 177


RESULT   14
ID   BAXD_HUMAN     STANDARD;      PRT;   143 AA.
AC   P55269;
DT   01-OCT-1996 (REL. 34, CREATED)
DT   01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   BAX PROTEIN, CYTOPLASMIC ISOFORM DELTA.
GN   BAX.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 95331797.
RA   APTE S.S., MATTEI M.-G., OLSEN B.R.;
RL   GENOMICS 26:592-594(1995).
CC   -!- ALTERNATIVE PRODUCTS: THE MEMBRANE ISOFORM ALPHA AND THE THREE
CC       CYTOPLASMIC ISOFORMS, BETA, GAMMA AND DELTA ARE GENERATED BY
CC       ALTERNATIVE SPLICING.
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC (POTENTIAL).
CC   -!- SIMILARITY: BELONGS TO THE BCL-2 FAMILY.
DR   EMBL; U19599; G841238; -.
DR   MIM; 600040; -.
KW   APOPTOSIS; ALTERNATIVE SPLICING.
SQ   SEQUENCE   143 AA;  15772 MW;  71AA1CBD CRC32;

  Query Match           11.3%;  Score 177;  DB 1;  Length 143;
  Best Local Similarity  27.4%;  Pred. No. 9.88e-13;
  Matches   29;  Conservative   31;  Mismatches  43;  Indels   3;  Gaps   3;

Db     36 tdsprevffrvaadmfsdgnfnwgrvvalfyfasklvlkalctkvpelirtimgwtldf- 94
          :::: | | ::|: :|  | :|||||||: |: :| |      :  ::  :    :||
Qy    104 AENAYEYFTKIATSLFESG-INWGRVVALLGFGYRLALHVYQHGLTGFLGQVTRFVVDFM 162

Db     95 lrerllgwiqdqggwdgllsyfgtptwqtvtifvagvltasltiwk 140
          |:  :  || ::||| : |:    |    | : :: || : : :::
Qy    163 LHHCIARWIAQRGGWVAALNLGNGPI-LNVLVVLGVVLLGQFVVRR 207


RESULT   15
ID   HSA1_MOUSE     STANDARD;      PRT;   172 AA.
AC   Q07440;
DT   01-FEB-1995 (REL. 31, CREATED)
DT   01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)
DT   01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)
DE   HEMOPOIETIC-SPECIFIC EARLY RESPONSE PROTEIN (A1 PROTEIN).
GN   A1.
OS   MUS MUSCULUS (MOUSE).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; RODENTIA.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CBA/J; TISSUE=BONE MARROW;
RX   MEDLINE; 93346743.
RA   LIN E.Y., ORLOFSKY A., BERGER M.S., PRYSTOWSKY M.B.;
RL   J. IMMUNOL. 151:1979-1988(1993).
CC   -!- FUNCTION: MAY FUNCTION IN THE RESPONSE OF HEMOPOIETIC CELLS TO
CC       EXTERNAL SIGNALS.
CC   -!- INDUCTION: BY GRANULOCYTE-MACROPHAGE COLONY-STIMULATING FACTOR
CC       AND LPS IN MACROPHAGES.
CC   -!- SUBCELLULAR LOCATION: INTRACELLULAR.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN HEMOPOIETIC TISSUES, INCLUDING
CC       BONE MARROW, SPLEEN AND THYMUS.
CC   -!- SIMILARITY: BELONGS TO THE BCL-2 FAMILY.
DR   EMBL; L16462; G293274; -.
DR   PROSITE; PS01080; BCL2.
KW   APOPTOSIS.
FT   DOMAIN       24     33       ALA/PRO-RICH.
SQ   SEQUENCE   172 AA;  19914 MW;  FFD38D6F CRC32;

  Query Match           11.3%;  Score 176;  DB 5;  Length 172;
  Best Local Similarity  28.9%;  Pred. No. 1.46e-12;
  Matches   33;  Conservative   29;  Mismatches  46;  Indels   6;  Gaps   6;

Db     28 sapsqacrvlqrvafsvqkeveknlksylddfhvesidtariifnqvmekefedgiinwg 87
          |: :|  | |  ::  : :  : :: : |: ::  : : |   |  :    || || |||
Qy     69 STMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQ-PTAENAYEYFTKIATSLFESGI-NWG 126

Db     88 rivtifafggvllkklpqeqialdvcaykqvssfvaefimnn-tgewirqnggw 140
          |:|::::||   |  |  | :|     ||: || :|:::::  : || | |||
Qy    127 RVVALLGFG-YRLA-LHVYQHGLTG-FLGQVTRFVVDFMLHHCIARWIAQRGGW 177
```

Search completed: Wed Aug 20 11:02:15 1997
Job time : 21 secs.

```
*************************************************************************
MPSRCH (TM)
*************************************************************************
```



```
MPsrch_nn   n.a. -  n.a. database search, using Smith-Waterman algorithm

Run on:          Thu Aug 21 12:30:31 1997;  MasPar time 1065.82 Seconds
                                             1326.187 Million cell updates/sec
Tabular output not generated.

Title:           >US-08-320-157-8
Description:     (1-1287) from US08320157.seq
Perfect Score:   1287
N.A. Sequence:      1 TTTTAATATAAATTAATGTG..........CCTCAAGAGTACAGAAGCTT 1287
     Comp:            AAAATTATATTTAATTACAC..........GGAGTTCTCATGTCTTCGAA

Scoring table:   TABLE default
                 Gap 6

Nmatch   STD :   Dbase 0;  Query 0

Searched:        362067 seqs, 549138275 bases x 2

Post-processing: Minimum Match 0%
                 Listing first 45 summaries

Database:        embl-new3
                   1:BCT 2:FUN 3:GEN1 4:GEN2 5:HTG1 6:HTG2 7:HUM 8:INV
                   9:ORG 10:MAM 11:VRT 12:PLN 13:PRO 14:ROD 15:SYN 16:UNC
                   17:VIR
Database:        genbank99
                   18:BCT1 19:BCT2 20:BCT3 21:BCT4 22:BCT5 23:BCT6 24:BCT7
                   25:BCT8 26:BCT9 27:BCT10 28:BCT11 29:GEN1 30:GEN2
                   31:GEN3 32:HTG1 33:HTG2 34:HTG3 35:INV1 36:INV2 37:INV3
                   38:INV4 39:INV5 40:INV6 41:INV7 42:INV8 43:INV9 44:INV10
                   45:INV11 46:MAM1 47:MAM2 48:MAM3 49:VRT1 50:VRT2 51:VRT3
                   52:VRT4 53:PAT1 54:PAT2 55:PAT3 56:PAT4 57:PAT5 58:PHG
                   59:PLN1 60:PLN2 61:PLN3 62:PLN4 63:PLN5 64:PLN6 65:PLN7
                   66:PLN8 67:PLN9 68:PLN10 69:PLN11 70:PRI1 71:PRI2
                   72:PRI3 73:PRI4 74:PRI5 75:PRI6 76:PRI7 77:PRI8 78:PRI9
                   79:PRI10 80:PRI11 81:PRI12 82:PRI13 83:PRI14 84:PRI15
                   85:ROD1 86:ROD2 87:ROD3 88:ROD4 89:ROD5 90:ROD6 91:ROD7
                   92:ROD8 93:STR 94:SYN 95:UNA 96:VRL1 97:VRL2 98:VRL3
                   99:VRL4 100:VRL5 101:VRL6 102:VRL7 103:VRL8 104:VRL9
                   105:VRL10
Database:        genbank-new3
                   106:BCT 107:GEN1 108:GEN2 109:HTG1 110:HTG2 111:INV
                   112:MAM 113:VRT 114:PHG 115:PLN 116:PRI1 117:PRI2
                   118:ROD 119:SYN 120:UNA 121:VRL
Database:        u-embl50_99
                   122:part1

Statistics:      Mean 11.725;  Variance 5.132;  scale 2.285
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
	1	1269	98.6	6478	77	HSU16812	Human Bak-2 gene, com	0.00e+00
	2	894	69.5	1360	73	HSCEBP1	H.sapiens BAK mRNA fo	0.00e+00
	3	890	69.2	2094	77	HSU16811	Human Bak mRNA, compl	0.00e+00
	4	768	59.7	1949	77	HSU23765	Human Bak protein mRN	0.00e+00
	5	753	58.5	5408	77	HSU16813	Human Bak-3 pseudogen	0.00e+00
	6	198	15.4	123579	5	HS291J10	Human DNA sequence **	3.97e-149
	7	177	13.8	444	70	D88396S2	Human DNA for apotosi	1.52e-129
	8	88	6.8	140	70	D88396S1	Human DNA for apotosi	4.96e-49
	9	37	2.9	215	57	I28278	Sequence 5 from paten	4.92e-08
c	10	35	2.7	215	57	I28278	Sequence 5 from paten	1.16e-06
c	11	27	2.1	250	82	HUMPECAM16	Homo sapiens platelet	1.44e-01
	12	27	2.1	1049	45	TTG8SCRNA	T.thermophila G8-scRN	1.44e-01
	13	27	2.1	2648	44	PFTUBAI	Plasmodium falciparum	1.44e-01
c	14	27	2.1	9747	52	XLPOMCA	X.laevis POMC-A gene	1.44e-01
c	15	26	2.0	610	45	XPU65216	Xenos pecki 28S ribos	5.48e-01
	16	26	2.0	2006	22	EFU39859	Enterococcus faecalis	5.48e-01
	17	26	2.0	2759	43	MILTMXC1	Leishmania tarentolae	5.48e-01
	18	26	2.0	20992	42	LEIKPMAX	Leishmania tarentolae	5.48e-01
c	19	26	2.0	91017	117	HS162C6	Human DNA sequence fr	5.48e-01
c	20	26	2.0	91017	7	HS162C6	Human DNA sequence fr	5.48e-01
c	21	26	2.0	91022	33	HS162C6	Human DNA sequence **	5.48e-01
	22	26	2.0	129207	34	HS363L9	Human DNA sequence **	5.48e-01
	23	24	1.9	1055	85	MMGASTGN3	M.musculus gast gene	7.12e+00
	24	24	1.9	1630	63	MTWSORFA	W.suaveolens mitochon	7.12e+00
c	25	25	1.9	1812	105	VSVGLYPN	Vesicular stomatitis	2.01e+00
c	26	25	1.9	1812	105	VSVGLYPZ	Vesicular stomatitis	2.01e+00
	27	24	1.9	1862	64	PSMANC	Piromyces sp. mRNA fo	7.12e+00
c	28	24	1.9	2220	39	DDCAMP1	D.discoideum CABP1 ge	7.12e+00
	29	24	1.9	2744	86	MMNMYC2	Woodchuck intronless	7.12e+00
	30	24	1.9	2857	74	HSE2F1S03	Human transcription f	7.12e+00
c	31	24	1.9	3474	45	TRBVSGH	Trypanosoma brucei MV	7.12e+00
	32	24	1.9	4484	61	CRSTR1	C.roseus str1 gene.	7.12e+00
c	33	25	1.9	5745	20	CPITOXIAB	C.perfringens DNA for	2.01e+00
c	34	24	1.9	7792	41	DROFSHA	D.melanogaster fsh me	7.12e+00
	35	25	1.9	7849	67	SOYPHYB	Soybean phytochrome B	2.01e+00
c	36	24	1.9	17279	52	TFLGTFTBPA	Snake (habu) gTfTBP g	7.12e+00
c	37	25	1.9	17366	52	TRUGTGTBPB	Snake (green habu) gT	2.01e+00
c	38	24	1.9	29203	37	CELF55E10	Caenorhabditis elegan	7.12e+00
c	39	24	1.9	34863	78	HSU73649	Human chromsome 11 97	7.12e+00
c	40	24	1.9	35609	73	HSCU90B3	Human DNA sequence fr	7.12e+00
c	41	24	1.9	37136	122	HSSRL9A13	Human DNA sequence fr	7.12e+00
c	42	24	1.9	37136	117	HSSRL9A13	Human DNA sequence fr	7.12e+00
c	43	24	1.9	96355	71	HS2A2B	Human DNA sequence fr	7.12e+00
c	44	24	1.9	96355	117	HS2A2B	Human DNA sequence fr	7.12e+00
	45	24	1.9	110384	73	HSAC000117	Human BAC clone RG062	7.12e+00

ALIGNMENTS

```
RESULT      1
LOCUS       HSU16812      6478 bp    DNA             PRI       19-AUG-1995
DEFINITION  Human Bak-2 gene, complete cds.
ACCESSION   U16812
NID         g595925
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 6478)
  AUTHORS   Kiefer,M.C., Brauer,M.J., Powers,V.C., Wu,J.J., Umansky,S.R.,
            Tomei,L.D. and Barr,P.J.
  TITLE     Modulation of apoptosis by the widely distributed Bcl-2 homologue
            Bak
  JOURNAL   Nature 374 (6524), 736-739 (1995)
  MEDLINE   95231654
REFERENCE   2  (bases 1 to 6478)
  AUTHORS   Kiefer,M.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-NOV-1994) Michael C. Kiefer, Mol. Biol., LXR
```

```
          Biotechnology Inc., 1401 Marina Way South, Richmond, CA 94804, USA
FEATURES             Location/Qualifiers
     source          1..6478
                     /organism="Homo sapiens"
                     /clone="A4"
                     /clone_lib="pWE 15/Human genomic DNA"
                     /chromosome="20"
     CDS             3279..3914
                     /gene="Bak-2"
                     /codon_start=1
                     /product="Bak-2 protein"
                     /db_xref="PID:g595926"
                     /translation="MASGQGPGPPRQECGEPALPSASEEQVAQDTEEVFRSYVFYHHQ
                     QEQEAEGAAAPADPEMVTLPLQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQP
                     TAENAYEYFTKIASSLFESGINWGRVVALLGFSYRLALHIYQRGLTGFLGQVTRFVVD
                     FMLHHCIARWIAQRGGWVAALNLGNGPILNVLVVLGVVLLGQFVVRRFFKS"
BASE COUNT     1503 a    1609 c    1599 g    1767 t
ORIGIN

  Query Match           98.6%;  Score 1269;  DB 77;  Length 6478;
  Best Local Similarity  99.8%;  Pred. No. 0.00e+00;
  Matches  1287;  Conservative    0;  Mismatches   0;  Indels   3;  Gaps   3;

Db   2735 ttttaatataaattaatgtgctctatttatagagacaatacatgaaatatacttaataaa 2794
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      1 TTTTAATATAAATTAATGTGCTCTATTTATAGAGACAATACATGAAATATACTTAATAAA 60

Db   2795 aattcaaatgttatagaactgaaaaagatgaaaagtaaaaacaacctattccccagaggt 2854
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     61 AATTCAAATGTTATAGAACTGAAAAAGATGAAAAGTAAAAACAACCTATTCCCCAGAGGT 120

Db   2855 agccactgtccatagtttctattttagattctttcctttatacaagattattatagcttc 2914
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    121 AGCCACTGTCCATAGTTTCTATTTTAGATTCTTTCCTTTATACAAGATTATTATAGCTTC 180

Db   2915 tatttttttggtgtatgaactgtagtcctagaggattttattagttatgagttctataact 2974
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    181 TATTTTTTGGTGTATGAACTGTAGTCCTAGAGGATTTTATTAGTTATGAGTTCTATAACT 240

Db   2975 aagatccatcatcttagttgctaagaacgtagatactgagaacatcatttaaaaaaacat 3034
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    241 AAGATCCATCATCTTAGTTGCTAAGAACGTAGATACTGAGAACATCATTTAAAAAAACAT 300

Db   3035 ttttggctggcacctctatgatcactggagtctcgcgggtccctcaggctgcacagggac 3094
           |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Qy    301 TTTTGGCTGGCACCTC-ATGATCACTGGAGTCTCGCGGGTCCCTCAGGCTGCACAGGGAC 359

Db   3095 aagtaaaggctacatccagatgctgggaatgcactgacgcccattcctggaaactgggct 3154
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    360 AAGTAAAGGCTACATCCAGATGCTGGGAATGCACTGACGCCCATTCCTGGAAACTGGGCT 419

Db   3155 cccactcagcccctgggagcagcagccgccagcccctcgggacctccatctccaccctgc 3214
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    420 CCCACTCAGCCCCTGGGAGCAGCAGCCGCCAGCCCCTCGGGACCTCCATCTCCACCCTGC 479

Db   3215 tgagccacccgggttgggccaggatcccggcaggctgatcccgtcctccactgagacctg 3274
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    480 TGAGCCACCCGGGTTGGGCCAGGATCCCGGCAGGCTGATCCCGTCCTCCACTGAGACCTG 539

Db   3275 aaaaatggcttcggggcaaggcccaggtcctcccaggcaggagtgcggagagcctgccct 3334
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    540 AAAAATGGCTTCGGGGCAAGGCCCAGGTCCTCCCAGGCAGGAGTGCGGAGAGCCTGCCCT 599

Db   3335 gccctctgcttctgaggagcaggtagcccaggacacagaggaggttttccgcagctacgt 3394
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    600 GCCCTCTGCTTCTGAGGAGCAGGTAGCCCAGGACACAGAGGAGGTTTTCCGCAGCTACGT 659

Db   3395 tttttaccaccatcagcaggaacaggaggctgaaggggcggctgcccctgccgacccaga 3454
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    660 TTTTTACCACCATCAGCAGGAACAGGAGGCTGAAGGGGCGGCTGCCCCTGCCGACCCAGA 719

Db   3455 gatggtcaccttacctctgcaacctagcagcaccatggggcaggtgggacggcagctcgc 3514
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    720 GATGGTCACCTTACCTCTGCAACCTAGCAGCACCATGGGGCAGGTGGGACGGCAGCTCGC 779

Db   3515 catcattggggacgacatcaaccgacgctatgactcagagttccagaccatgttgcagca 3574
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    780 CATCATTGGGGACGACATCAACCGACGCTATGACTCAGAGTTCCAGACCATGTTGCAGCA 839

Db   3575 cctgcagcccacggcagagaatgcctatgagtacttcaccaagattgcctccagcctgtt 3634
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    840 CCTGCAGCCCACGGCAGAGAATGCCTATGAGTACTTCACCAAGATTGCCTCCAGCCTGTT 899

Db   3635 tgagagtggcatcaattggggccgtgtggtggctcttctgggcttcagctaccgtctggc 3694
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    900 TGAGAGTGGCATCAATTGGGGCCGTGTGGTGGCTCTTCTGGGCTTCAGCTACCGTCTGGC 959

Db   3695 cctacacatctaccagcgtggcctgactggcttcctgggccaggtgacccgctttgtggt 3754
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    960 CCTACACATCTACCAGCGTGGCCTGACTGGCTTCCTGGGCCAGGTGACCCGCTTTGTGGT 1019

Db   3755 ggacttcatgctgcatcactgcattgcccggtggattgcacagaggggtggctgggtggc 3814
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1020 GGACTTCATGCTGCATCACTGCATTGCCCGGTGGATTGCACAGAGGGGTGGCTGGGTGGC 1079

Db   3815 agccctgaacttgggcaatggtcccatcctgaacgtgctggtggttctgggtgtggttct 3874
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1080 AGCCCTGAACTTGGGCAATGGTCCCATCCTGAACGTGCTGGTGGTTCTGGGTGTGGTTCT 1139

Db   3875 gttgggccagtttgtggtacgaagattcttcaaatcatgactcccaagggtgccctttgg 3934
           |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Qy   1140 GTTGGGCCAGTTTGTGGTACGAAGATTCTTCAAATCATGACTCCCAAGGGTGCC-TTTGG 1198

Db   3935 ggtcccagttcagacccctgcctggacttaagcgaagtctttgccttctctgctccttgc 3994
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1199 GGTCCCAGTTCAGACCCCTGCCTGGACTTAAGCGAAGTCTTTGCCTTCTCTGCTCCTTGC 1258

Db   3995 aggggtccccccctcaagagtacagaagctt 4024
           |||| ||||||||||||||||||||||||||
Qy   1259 AGGG-TCCCCCCTCAAGAGTACAGAAGCTT 1287


RESULT    2
LOCUS       HSCEBP1      1360 bp    RNA             PRI       05-MAY-1995
DEFINITION  H.sapiens BAK mRNA for BCl-2 homologue.
ACCESSION   X84213
NID         g804984
KEYWORDS    Bcl-2 protein; CEBP-1 gene.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1360)
  AUTHORS   Farrow,S.N., White,J.H., Martinou,I., Raven,T., Pun,K.T.,
            Grinham,C.J., Martinou,J.C. and Brown,R.
  TITLE     Cloning of a bcl-2 homologue by interaction with adenovirus E1B 19K
  JOURNAL   Nature 374 (6524), 731-733 (1995)
  MEDLINE   95231652
  REMARK    Erratum:[Nature 1995 Jun 1;375(6530):431]]
REFERENCE   2  (bases 1 to 1360)
  AUTHORS   Brown,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-JAN-1995) R. Brown, Glaxo Research & Development,
            Greenford Road, Greenford, Middlesex UB6 0HE, UK
FEATURES             Location/Qualifiers
     source          1..1360
                     /organism="Homo sapiens"
                     /tissue_type="B-cell"
                     /clone_lib="EBV-transformed B-cell"
     CDS             193..828
                     /gene="BAK"
                     /codon_start=1
                     /product="bcl-2 homologue"
                     /db_xref="PID:g804985"
                     /translation="MASGQGPGPPRQECGEPALPSASEEQVAQDTEEVFRSYVFYRHQ
                     QEQEAEGVAAPADPEMVTLPLQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQP
```

```
                    TAENAYEYFTKIATSLFESGINWGRVVALLGFGYRLALHVYQHGLTGFLGQVTRFVVD
                    FMLHHCIARWIAQRGGWVAALNLGNGPILNVLVVLGVVLLGQFVVRRFFKS"
BASE COUNT      257 a    405 c    400 g    298 t
ORIGIN

  Query Match           69.5%;  Score 894;  DB 73;  Length 1360;
  Best Local Similarity  98.2%;  Pred. No. 0.00e+00;
  Matches   921;  Conservative    0;  Mismatches  15;  Indels   2;  Gaps   2;

Db      1 acagggacaagtaaaggctacatccagatgccgggaatgcactgacgcccattcctggaa 60
          |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Qy    352 ACAGGGACAAGTAAAGGCTACATCCAGATGCTGGGAATGCACTGACGCCCATTCCTGGAA 411

Db     61 actgggctcccactcagcccctgggagcagcagccgccagcccctcgggacctccatctc 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    412 ACTGGGCTCCCACTCAGCCCCTGGGAGCAGCAGCCGCCAGCCCCTCGGGACCTCCATCTC 471

Db    121 caccctgctgagccacccgggttgggccaggatcccggcaggctgatcccgtcctccact 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    472 CACCCTGCTGAGCCACCCGGGTTGGGCCAGGATCCCGGCAGGCTGATCCCGTCCTCCACT 531

Db    181 gagacctgaaaaatggcttcggggcaaggcccaggtcctcccaggcaggagtgcggagag 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    532 GAGACCTGAAAAATGGCTTCGGGGCAAGGCCCAGGTCCTCCCAGGCAGGAGTGCGGAGAG 591

Db    241 cctgccctgccctctgcttctgaggagcaggtagcccaggacacagaggaggttttccgc 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    592 CCTGCCCTGCCCTCTGCTTCTGAGGAGCAGGTAGCCCAGGACACAGAGGAGGTTTTCCGC 651

Db    301 agctacgtttttttaccgccatcagcaggaacaggaggctgaaggggtggctgcccctgcc 360
          |||||||||||||||| |||||||||||||||||||||||||||| ||||||||||||||
Qy    652 AGCTACGTTTTTTACCACCATCAGCAGGAACAGGAGGCTGAAGGGGCGGCTGCCCCTGCC 711

Db    361 gacccagagatggtcaccttacctctgcaacctagcagcaccatggggcaggtgggacgg 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    712 GACCCAGAGATGGTCACCTTACCTCTGCAACCTAGCAGCACCATGGGGCAGGTGGGACGG 771

Db    421 cagctcgccatcatcggggacgacatcaaccgacgctatgactcagagttccagaccatg 480
          |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Qy    772 CAGCTCGCCATCATTGGGGACGACATCAACCGACGCTATGACTCAGAGTTCCAGACCATG 831

Db    481 ttgcagcacctgcagcccacggcagagaatgcctatgagtacttcaccaagattgccacc 540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Qy    832 TTGCAGCACCTGCAGCCCACGGCAGAGAATGCCTATGAGTACTTCACCAAGATTGCCTCC 891

Db    541 agcctgtttgagagtggcatcaattggggccgtgtggtggctcttctgggcttcggctac 600
          |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Qy    892 AGCCTGTTTGAGAGTGGCATCAATTGGGGCCGTGTGGTGGCTCTTCTGGGCTTCAGCTAC 951

Db    601 cgtctggccctacacgtctaccagcatggcctgactggcttcctaggccaggtgacccgc 660
          |||||||||||||| ||||||||| |||||||||||||||||| ||||||||||||||||
Qy    952 CGTCTGGCCCTACACATCTACCAGCGTGGCCTGACTGGCTTCCTGGGCCAGGTGACCCGC 1011

Db    661 ttcgtggtcgacttcatgctgcatcactgcattgcccggtggattgcacagaggggtggc 720
          || ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1012 TTTGTGGTGGACTTCATGCTGCATCACTGCATTGCCCGGTGGATTGCACAGAGGGGTGGC 1071

Db    721 tgggtggcagccctgaacttgggcaatggtcccatcctgaacgtgctggtggttctgggt 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1072 TGGGTGGCAGCCCTGAACTTGGGCAATGGTCCCATCCTGAACGTGCTGGTGGTTCTGGGT 1131

Db    781 gtggttctgttgggccagtttgtggtacgaagattcttcaaatcatgactcccaagggtg 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1132 GTGGTTCTGTTGGGCCAGTTTGTGGTACGAAGATTCTTCAAATCATGACTCCCAAGGGTG 1191

Db    841 ccctttgggtcccggttcagacccctgcctggacttaagcgaagtctttgccttctctgt 900
          || || ||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Qy   1192 CCTTTGGGGTCCCAGTTCAGACCCCTGCCTGGACTTAAGCGAAGTCTTTGCCTTCTCTGC 1251

Db    901 tcccttgcaggggtccccccctcaagagtacagaagctt 938
          ||| |||||||| ||||||||||||||||||||||||||
Qy   1252 TCC-TTGCAGGG-TCCCCCCTCAAGAGTACAGAAGCTT 1287
```

```
RESULT    3
LOCUS       HSU16811     2094 bp    mRNA            PRI       19-AUG-1995
DEFINITION  Human Bak mRNA, complete cds.
ACCESSION   U16811
NID         g595923
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2094)
  AUTHORS   Kiefer,M.C., Brauer,M.J., Powers,V.C., Wu,J.J., Umansky,S.R.,
            Tomei,L.D. and Barr,P.J.
  TITLE     Modulation of apoptosis by the widely distributed Bcl-2 homologue
            Bak
  JOURNAL   Nature 374 (6524), 736-739 (1995)
  MEDLINE   95231654
REFERENCE   2  (bases 1 to 2094)
  AUTHORS   Kiefer,M.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-NOV-1994) Michael C. Kiefer, Mol. Biol., LXR
            Biotechnology Inc., 1401 Marina Way South, Richmond, CA 94804, USA
FEATURES             Location/Qualifiers
     source          1..2094
                     /organism="Homo sapiens"
                     /clone="W5"
                     /clone_lib="WI-L2/ZAP II"
                     /chromosome="6"
                     /map="6p21-23"
                     /cell_line="EBV transformed B lymphoblastoid"
     CDS             201..836
                     /gene="Bak"
                     /codon_start=1
                     /product="Bak protein"
                     /db_xref="PID:g595924"
                     /translation="MASGQGPGPPRQECGEPALPSASEEQVAQDTEEVFRSYVFYRHQ
                     QEQEAEGVAAPADPEMVTLPLQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQP
                     TAENAYEYFTKIATSLFESGINWGRVVALLGFGYRLALHVYQHGLTGFLGQVTRFVVD
                     FMLHHCIARWIAQRGGWVAALNLGNGPILNVLVVLGVVLLGQFVVRRFFKS"
BASE COUNT      410 a    608 c    606 g    470 t
ORIGIN

  Query Match           69.2%;  Score 890;  DB 77;  Length 2094;
  Best Local Similarity  98.2%;  Pred. No. 0.00e+00;
  Matches   917;  Conservative    0;  Mismatches  15;  Indels   2;  Gaps   2;

Db     13 gggacaagtaaaggctacatccagatgccgggaatgcactgacgcccattcctggaaact 72
          |||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Qy    355 GGGACAAGTAAAGGCTACATCCAGATGCTGGGAATGCACTGACGCCCATTCCTGGAAACT 414

Db     73 gggctcccactcagcccctgggagcagcagccgccagcccctcgg-acctccatctccac 131
          |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Qy    415 GGGCTCCCACTCAGCCCCTGGGAGCAGCAGCCGCCAGCCCCTCGGGACCTCCATCTCCAC 474

Db    132 cctgctgagccacccgggttgggccaggatcccggcaggctgatcccgtcctccactgag 191
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    475 CCTGCTGAGCCACCCGGGTTGGGCCAGGATCCCGGCAGGCTGATCCCGTCCTCCACTGAG 534

Db    192 acctgaaaaatggcttcggggcaaggcccaggtcctcccaggcaggagtgcggagagcct 251
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    535 ACCTGAAAAATGGCTTCGGGGCAAGGCCCAGGTCCTCCCAGGCAGGAGTGCGGAGAGCCT 594

Db    252 gccctgccctctgcttctgaggagcaggtagcccaggacacagaggaggttttccgcagc 311
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    595 GCCCTGCCCTCTGCTTCTGAGGAGCAGGTAGCCCAGGACACAGAGGAGGTTTTCCGCAGC 654

Db    312 tacgtttttttaccgccatcagcaggaacaggaggctgaaggggtggctgcccctgccgac 371
          |||||||||||||| |||||||||||||||||||||||||||| ||||||||||||||||
Qy    655 TACGTTTTTTACCACCATCAGCAGGAACAGGAGGCTGAAGGGGCGGCTGCCCCTGCCGAC 714

Db    372 ccagagatggtcaccttacctctgcaacctagcagcaccatggggcaggtgggacggcag 431
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Qy    715 CCAGAGATGGTCACCTTACCTCTGCAACCTAGCAGCACCATGGGGCAGGTGGGACGGCAG 774

Db    432 ctcgccatcatcggggacgacatcaaccgacgctatgactcagagttccagaccatgttg 491
          |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Qy    775 CTCGCCATCATTGGGGACGACATCAACCGACGCTATGACTCAGAGTTCCAGACCATGTTG 834

Db    492 cagcacctgcagcccacggcagagaatgcctatgagtacttcaccaagattgccaccagc 551
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Qy    835 CAGCACCTGCAGCCCACGGCAGAGAATGCCTATGAGTACTTCACCAAGATTGCCTCCAGC 894

Db    552 ctgtttgagagtggcatcaattggggccgtgtggtggctcttctgggcttcggctaccgt 611
          |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Qy    895 CTGTTTGAGAGTGGCATCAATTGGGGCCGTGTGGTGGCTCTTCTGGGCTTCAGCTACCGT 954

Db    612 ctggccctacacgtctaccagcatggcctgactggcttcctaggccaggtgacccgcttc 671
          |||||||||||| ||||||||| |||||||||||||||||| ||||||||||||||||||
Qy    955 CTGGCCCTACACATCTACCAGCGTGGCCTGACTGGCTTCCTGGGCCAGGTGACCCGCTTT 1014

Db    672 gtggtcgacttcatgctgcatcactgcattgcccggtggattgcacagaggggtggctgg 731
          |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1015 GTGGTGGACTTCATGCTGCATCACTGCATTGCCCGGTGGATTGCACAGAGGGGTGGCTGG 1074

Db    732 gtggcagccctgaacttgggcaatggtcccatcctgaacgtgctggtggttctgggtgtg 791
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1075 GTGGCAGCCCTGAACTTGGGCAATGGTCCCATCCTGAACGTGCTGGTGGTTCTGGGTGTG 1134

Db    792 gttctgttgggccagtttgtggtacgaagattcttcaaatcatgactcccaagggtgccc 851
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1135 GTTCTGTTGGGCCAGTTTGTGGTACGAAGATTCTTCAAATCATGACTCCCAAGGGTGCCT 1194

Db    852 tttgggtcccggttcagacccctgcctggacttaagcgaagtctttgccttctctgttcc 911
          || ||||||| |||||||||||||||||||||||||||||||||||||||||||||| |||
Qy   1195 TTGGGGTCCCAGTTCAGACCCCTGCCTGGACTTAAGCGAAGTCTTTGCCTTCTCTGCTCC 1254

Db    912 cttgcagggtccccccctcaagagtacagaagctt 945
          |||||||||||||||||||||||||||||||||
Qy   1255 -TTGCAGGGTCCCCCCTCAAGAGTACAGAAGCTT 1287


RESULT    4
LOCUS       HSU23765     1949 bp    mRNA            PRI       29-MAR-1996
DEFINITION  Human Bak protein mRNA, complete cds.
ACCESSION   U23765
NID         g758797
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1949)
  AUTHORS   Chittenden,T., Harrington,E.A., O'Connor,R., Flemington,C.,
            Lutz,R.J., Evan,G.I. and Guild,B.C.
  TITLE     Induction of apoptosis by the Bcl-2 homologue Bak
  JOURNAL   Nature 374 (6524), 733-736 (1995)
  MEDLINE   95231653
REFERENCE   2  (bases 1 to 1949)
  AUTHORS   Chittenden,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-MAR-1995) Thomas Chittenden, Apoptosis Technology
            Inc., 148 Sidney St., Cambridge, MA 02139, USA
FEATURES             Location/Qualifiers
     source          1..1949
                     /organism="Homo sapiens"
                     /clone_lib="Jurkat cell cDNA library (Stratagene)"
                     /cell_line="Jurkat cell"
     CDS             65..700
                     /function="induction of apoptosis"
                     /codon_start=1
                     /product="Bak protein"
                     /db_xref="PID:g758798"
                     /translation="MASGQGPGPPRQECGEPALPSASEEQVAQDTEEVFRSYVFYRHQ
                     QEQEAEGVAAPADPEMVTLPLQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQP
                     TAENAYEYFTKIATSLFESGINWGRVVALLGFGYRLALHVYQHGLTGFLGQVTRFVVD
                     FMLHHCIARWIAQRGGWVAALNLGNGPILNVLVVLGVVLLGQFVVRRFFKS"
     polyA_site      1949
BASE COUNT      368 a    557 c    575 g    449 t
ORIGIN

  Query Match           59.7%;  Score 768;  DB 77;  Length 1949;
  Best Local Similarity  98.0%;  Pred. No. 0.00e+00;
  Matches   794;  Conservative    0;  Mismatches  14;  Indels   2;  Gaps   2;

Db      1 tgagccacccgggttgggccaggatcccggcaggctgatcccgtcctccactgagacctg 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    480 TGAGCCACCCGGGTTGGGCCAGGATCCCGGCAGGCTGATCCCGTCCTCCACTGAGACCTG 539

Db     61 aaaaatggcttcggggcaaggcccaggtcctcccaggcaggagtgcggagagcctgccct 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    540 AAAAATGGCTTCGGGGCAAGGCCCAGGTCCTCCCAGGCAGGAGTGCGGAGAGCCTGCCCT 599

Db    121 gccctctgcttctgaggagcaggtagcccaggacacagaggaggttttccgcagctacgt 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    600 GCCCTCTGCTTCTGAGGAGCAGGTAGCCCAGGACACAGAGGAGGTTTTCCGCAGCTACGT 659

Db    181 tttttaccgccatcagcaggaacaggaggctgaaggggtggctgcccctgccgacccaga 240
          |||||||| |||||||||||||||||||||||||||||||| ||||||||||||||||||
Qy    660 TTTTTACCACCATCAGCAGGAACAGGAGGCTGAAGGGGCGGCTGCCCCTGCCGACCCAGA 719

Db    241 gatggtcaccttacctctgcaacctagcagcaccatggggcaggtgggacggcagctcgc 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    720 GATGGTCACCTTACCTCTGCAACCTAGCAGCACCATGGGGCAGGTGGGACGGCAGCTCGC 779

Db    301 catcatcggggacgacatcaaccgacgctatgactcagagttccagaccatgttgcagca 360
          |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    780 CATCATTGGGGACGACATCAACCGACGCTATGACTCAGAGTTCCAGACCATGTTGCAGCA 839

Db    361 cctgcagcccacggcagagaatgcctatgagtacttcaccaagattgccaccagcctgtt 420
          ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Qy    840 CCTGCAGCCCACGGCAGAGAATGCCTATGAGTACTTCACCAAGATTGCCTCCAGCCTGTT 899

Db    421 tgagagtggcatcaattggggccgtgtggtggctcttctgggcttcggctaccgtctggc 480
          |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Qy    900 TGAGAGTGGCATCAATTGGGGCCGTGTGGTGGCTCTTCTGGGCTTCAGCTACCGTCTGGC 959

Db    481 cctacacgtctaccagcatggcctgactggcttcctaggccaggtgacccgcttcgtggt 540
          ||||||| ||||||||| |||||||||||||||||| |||||||||||||||| |||||
Qy    960 CCTACACATCTACCAGCGTGGCCTGACTGGCTTCCTGGGCCAGGTGACCCGCTTTGTGGT 1019

Db    541 cgacttcatgctgcatcactgcattgcccggtggattgcacagaggggtggctgggtggc 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1020 GGACTTCATGCTGCATCACTGCATTGCCCGGTGGATTGCACAGAGGGGTGGCTGGGTGGC 1079

Db    601 agccctgaacttgggcaatggtcccatcctgaacgtgctggtggttctgggtgtggttct 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1080 AGCCCTGAACTTGGGCAATGGTCCCATCCTGAACGTGCTGGTGGTTCTGGGTGTGGTTCT 1139

Db    661 gttgggccagtttgtggtacgaagattcttcaaatcatgactcccaagggtgccctttgg 720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||| || ||
Qy   1140 GTTGGGCCAGTTTGTGGTACGAAGATTCTTCAAATCATGACTCCCAAGGGTGCCTTTGGG 1199

Db    721 gtcccggttcagacccctgcctggacttaagcgaagtctttgccttctctgttcccttgc 780
          ||||| |||||||||||||||||||||||||||||||||||||||||||| ||| ||||
Qy   1200 GTCCCAGTTCAGACCCCTGCCTGGACTTAAGCGAAGTCTTTGCCTTCTCTGCTCC-TTGC 1258

Db    781 aggggtccccccctcaagagtacagaagctt 810
          ||||| ||||||||||||||||||||||||||
Qy   1259 AGGG-TCCCCCCTCAAGAGTACAGAAGCTT 1287


RESULT    5
LOCUS       HSU16813     5408 bp    DNA             PRI       19-AUG-1995
DEFINITION  Human Bak-3 pseudogene, complete cds.
ACCESSION   U16813
NID         g595927
KEYWORDS    .
SOURCE      human.
```

```
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 5408)
  AUTHORS   Kiefer,M.C., Brauer,M.J., Powers,V.C., Wu,J.J., Umansky,S.R.,
            Tomei,L.D. and Barr,P.J.
  TITLE     Modulation of apoptosis by the widely distributed Bcl-2 homologue
            Bak
  JOURNAL   Nature 374 (6524), 736-739 (1995)
  MEDLINE   95231654
REFERENCE   2  (bases 1 to 5408)
  AUTHORS   Kiefer,M.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-NOV-1994) Michael C. Kiefer, Mol. Biol., LXR
            Biotechnology Inc., 1401 Marina Way South, Richmond, CA 94804, USA
FEATURES             Location/Qualifiers
     source          1..5408
                     /organism="Homo sapiens"
                     /clone="C5"
                     /clone_lib="pWE 15/Human genomic DNA"
                     /chromosome="11"
     CDS             1665..1931
                     /gene="Bak-3"
                     /pseudo
BASE COUNT     1369 a   1384 c   1314 g   1341 t
ORIGIN

  Query Match           58.5%;  Score 753;  DB 77;  Length 5408;
  Best Local Similarity  91.0%;  Pred. No. 0.00e+00;
  Matches   886;  Conservative    0;  Mismatches  79;  Indels   9;  Gaps   7;

Db   1438 atgatcactggagtctcacgggtccctcgggccacacagggacaagcaaaggctacatcc 1497
            |||||||||||||||| |||||||||| |||  |||||||||||| ||||||||||||
Qy    317 ATGATCACTGGAGTCTCGCGGGTCCCTCAGGCTGCACAGGGACAAGTAAAGGCTACATCC 376

Db   1498 agataccagaaatgcactgacgcccgttcctggaagctgggctcccactcagcccctggg 1557
            |||| |  | ||||||||||||||| |||||||| |||||||||||||||||||||||
Qy    377 AGATGCTGGGAATGCACTGACGCCCATTCCTGGAAACTGGGCTCCCACTCAGCCCCTGGG 436

Db   1558 agcagcagcctccagccccttgggaccttcaactccaccctgctgacccacgcgggttga 1617
            |||||||||| |||||||| ||||||| || |||||||||||||| |||| |||||||
Qy    437 AGCAGCAGCCGCCAGCCCCTCGGGACCTCCATCTCCACCCTGCTGAGCCACCCGGGTTGG 496

Db   1618 gccagcatccctggaggctgacactgtcctccactgagacctgaaaaatggcatcggggc 1677
            ||||| ||||| | ||||||| | ||||||||||||||||||||||||||||| |||||||
Qy    497 GCCAGGATCCCGGCAGGCTGATCCCGTCCTCCACTGAGACCTGAAAAATGGCTTCGGGGC 556

Db   1678 aaggcccagggcctcccaggcaggagtgcggaaagcctgccctgccctctgcttctgagg 1737
            |||||||||| |||||||||||||||||||||||| |||||||||||||||||||||||||
Qy    557 AAGGCCCAGGTCCTCCCAGGCAGGAGTGCGGAGAGCCTGCCCTGCCCTCTGCTTCTGAGG 616

Db   1738 agcaggtagcccaggacatggagg-ggtttttccgcagctacgttttttaccaccatcagc 1796
            |||||||||||||||||  |||| ||||||||||||||||||||||||||||||||||
Qy    617 AGCAGGTAGCCCAGGACACAGAGGAGGTTTTCCGCAGCTACGTTTTTTACCACCATCAGC 676

Db   1797 aggaacaggaggctgaaggggcggccgcccctgccgacccagagatggtcaccttgcccc 1856
            |||||||||||||||||||||||| |||||||||||||||||||||||||||||| || |
Qy    677 AGGAACAGGAGGCTGAAGGGGCGGCTGCCCCTGCCGACCCAGAGATGGTCACCTTACCTC 736

Db   1857 tccaacctagcagcaccatggggcaggtgggacggcagctcgccatcaccagg-acgaca 1915
            | |||||||||||||||||||||||||||||||||||||||||||||||   || ||||||
Qy    737 TGCAACCTAGCAGCACCATGGGGCAGGTGGGACGGCAGCTCGCCATCATTGGGGACGACA 796

Db   1916 tcaaccggcactatgacttcggagttccagaccatgctgcagcacctgcagcccacggca 1975
            ||||||| | |||||||| | |||||||||||||||| ||||||||||||||||||||||
Qy    797 TCAACCGACGCTATGACT-CAGAGTTCCAGACCATGTTGCAGCACCTGCAGCCCACGGCA 855

Db   1976 gagaacgcctacgagtacttcaccaagatcgcctccagcctgtttgagagtggcatcaac 2035
            ||||| ||||| |||||||||||||||| |||||||||||||||||||||||||||||
Qy    856 GAGAATGCCTATGAGTACTTCACCAAGATTGCCTCCAGCCTGTTTGAGAGTGGCATCAAT 915

Db   2036 cggggccgtgtggtggctctcctgggcttcggctaccgtctggtcctacatgtctaccag 2095
            |||||||||||||||||| ||||||||| |||||||||||| ||||||  ||||||||

Qy    916 TGGGGCCGTGTGGTGGCTCTTCTGGGCTTCAGCTACCGTCTGGCCCTACACATCTACCAG 975

Db   2096 cacggcttgactggcttcctgggcctggtgacccgcttcgtggt---cttcatgctgcaa 2152
            |  ||| |||||||||||||||| |||||||||||| |||||   ||||||||||||
Qy    976 CGTGGCCTGACTGGCTTCCTGGGCCAGGTGACCCGCTTTGTGGTGGACTTCATGCTGCAT 1035

Db   2153 caaggcatcgcccggtggatctcgcagaggggcggctgggtggcagccctggacttgggc 2212
            ||  |||| ||||||||||||  | |||||||| |||||||||||||||||| ||||||||
Qy   1036 CACTGCATTGCCCGGTGGATTGCACAGAGGGGTGGCTGGGTGGCAGCCCTGAACTTGGGC 1095

Db   2213 aatagtcccatcctgaacgtgctggtggttgtgggtgtggttctgctgggccagtttgtg 2272
            ||| |||||||||||||||||||||||| ||||||||||||| ||||||||||||||
Qy   1096 AATGGTCCCATCCTGAACGTGCTGGTGGTTCTGGGTGTGGTTCTGTTGGGCCAGTTTGTG 1155

Db   2273 gtaagaagattcttcaaatcatgactcccaggggtgtcctttggggtcccagctgtgacc 2332
            ||| |||||||||||||||||||||||| ||||| |||||||||||||| |  ||||
Qy   1156 GTACGAAGATTCTTCAAATCATGACTCCCAAGGGTG-CCTTTGGGGTCCCAGTTCAGACC 1214

Db   2333 cctgcctggacttaagccaagtctttgccttccccactcccttgcaggggtcaccccttca 2392
            |||||||||||||||| |||||||||||||| |  |||| |||||||| || ||| |||
Qy   1215 CCTGCCTGGACTTAAGCGAAGTCTTTGCCTTCTCTGCTCC-TTGCAGGG-TCCCCCCTCA 1272

Db   2393 aaagtacagaagct 2406
            | ||||||||||||
Qy   1273 AGAGTACAGAAGCT 1286


RESULT    6
ID   HS291J10   standard; DNA; HTG; 123579 BP.
AC   Z93017;
NI   e1041698
DT   19-MAR-1997 (Rel. 51, Created)
DT   19-MAR-1997 (Rel. 51, Last updated, Version 1)
DE   Human DNA sequence *** SEQUENCING IN PROGRESS *** from clone 291J10
KW   HTG; HTGS_PHASE1.
OS   Homo sapiens (human)
OC   Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
OC   Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   1-123579
RA   Beck S.;
RT   ;
RL   Submitted (06-NOV-1996) to the EMBL/GenBank/DDBJ databases.
RL   Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CB10 1SA,
RL   UK. E-mail enquires: humquery@sanger.ac.uk Clone requests:
RL   clonerequest@sanger.ac.uk
CC   IMPORTANT: This sequence is unfinished and does not necessarily
CC   represent the correct sequence. Work on the sequence is in progress
CC   and
CC   the release of this data is based on the understanding that the
CC   sequence may change as work continues. The sequence may be
CC   contaminated
CC   with foreign sequence from E.coli, yeast, vector, phage etc.
CC   Order of segments is not known; 800 n's separate segments.
CC   Unfinished sequence: dJ291J10  Contig_ID: 01371  Length: 3951 bp
CC   Unfinished sequence: dJ291J10  Contig_ID: 01622  Length: 1040 bp
CC   Unfinished sequence: dJ291J10  Contig_ID: 00751  Length: 6361 bp
CC   Unfinished sequence: dJ291J10  Contig_ID: 01422  Length: 1722 bp
CC   Unfinished sequence: dJ291J10  Contig_ID: 00105  Length: 26804 bp
CC   Unfinished sequence: dJ291J10  Contig_ID: 01527  Length: 1598 bp
CC   Unfinished sequence: dJ291J10  Contig_ID: 01231  Length: 1547 bp
CC   Unfinished sequence: dJ291J10  Contig_ID: 01408  Length: 2194 bp
CC   Unfinished sequence: dJ291J10  Contig_ID: 00133  Length: 4250 bp
CC   Unfinished sequence: dJ291J10  Contig_ID: 01390  Length: 4036 bp
CC   Unfinished sequence: dJ291J10  Contig_ID: 01463  Length: 7484 bp
CC   Unfinished sequence: dJ291J10  Contig_ID: 00255  Length: 14415 bp
CC   Unfinished sequence: dJ291J10  Contig_ID: 01327  Length: 9377 bp
CC   Unfinished sequence: dJ291J10  Contig_ID: 01441  Length: 6493 bp
CC   Unfinished sequence: dJ291J10  Contig_ID: 00931  Length: 21107 bp
FH   Key             Location/Qualifiers
FH
FT   source          1..123579
FT                   /organism="Homo sapiens"
```

```
FT                   /clone="291J10"
FT                   /chromosome="6"
SQ   Sequence 123579 BP; 28765 A; 27932 C; 27951 G; 27528 T; 11403 other;

  Query Match           15.4%;  Score 198;  DB 5;  Length 123579;
  Best Local Similarity  98.6%;  Pred. No. 3.97e-149;
  Matches   206;  Conservative    0;  Mismatches   2;  Indels   1;  Gaps   1;

Db 107852 gctggcacctctatgatcactggagtctcgcgggtccctcgggctgcacagggacaagta 107911
          |||||||||| |||||||||||||||||||||||||| ||||||||||||||||||
Qy    306 GCTGGCACCTC-ATGATCACTGGAGTCTCGCGGGTCCCTCAGGCTGCACAGGGACAAGTA 364

Db 107912 aaggctacatccagatgccgggaatgcactgacgcccattcctggaaactgggctcccac 107971
          |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Qy    365 AAGGCTACATCCAGATGCTGGGAATGCACTGACGCCCATTCCTGGAAACTGGGCTCCCAC 424

Db 107972 tcagcccctgggagcagcagccgccagcccctcgggacctccatctccaccctgctgagc 108031
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    425 TCAGCCCCTGGGAGCAGCAGCCGCCAGCCCCTCGGGACCTCCATCTCCACCCTGCTGAGC 484

Db 108032 cacccgggttgggccaggatcccggcagg 108060
          |||||||||||||||||||||||||||||
Qy    485 CACCCGGGTTGGGCCAGGATCCCGGCAGG 513


RESULT    7
LOCUS       D88396S2      444 bp    DNA             PRI       10-JAN-1997
DEFINITION  Human DNA for apotosis-regulator Bak, exon 2, 3 and partial cds.
ACCESSION   D88397
NID         g1655492
KEYWORDS    BaK; apotosis-regulator.
SEGMENT     2 of 2
SOURCE      Homo sapiens DNA, clone:pGEMBak01-02.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 444)
  AUTHORS   Eguchi,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-OCT-1996) to the DDBJ/EMBL/GenBank databases.
            Hidetaka Eguchi, Saitama Cancer Center Research Institute,
            Depertment of Biochemistry; 818 Komuro, Ina, Kita-adachi-gun,
            Saitama 362, Japan (E-mail:hide@saitama-cc.go.jp,
            Tel:048-722-1111(ex.255), Fax:048-722-1739)
REFERENCE   2  (bases 1 to 444)
  AUTHORS   Eguchi,H. and Hayashi,S.
  TITLE     Estrogen alters expression of apoptosis-regulators, Bcl-2, Bcl-xL
            and Bak, as well as susceptibility to therapeutic agents of human
            breast cancer cells
  JOURNAL   Unpublished (1996)
FEATURES             Location/Qualifiers
     source          1..444
                     /organism="Homo sapiens"
                     /clone="pGEMBak01-02"
     CDS             join(D88396:26..90,51..231,358..>444)
                     /codon_start=1
                     /product="BaK"
                     /db_xref="PID:g1655494"
                     /translation="MLQHLQPTAENAYEYFTKIATSLFESGINWGRVVALLGFGYRLA
                     LHVYQHGLTGFLGQVTRFVVDFMLHHCIARWIAQRGGWVAALNLGNGPILNVLVVLGV
                     VLLGQFVVR"
     intron          1..50
                     /number=1
     exon            51..231
                     /number=2
     intron          232..357
                     /number=2
     exon            358..444
                     /number=3
BASE COUNT       58 a    133 c    132 g    119 t      2 others
ORIGIN

  Query Match           13.8%;  Score 177;  DB 70;  Length 444;
```

```
  Best Local Similarity  95.9%;  Pred. No. 1.52e-129;
  Matches   185;  Conservative    0;  Mismatches   8;  Indels   0;  Gaps   0;

Db     42 tgcccacagcctgtttgagagtggcatcaattggggccgtgtggtggctcttctgggctt 101
          ||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    885 TGCCTCCAGCCTGTTTGAGAGTGGCATCAATTGGGGCCGTGTGGTGGCTCTTCTGGGCTT 944

Db    102 cggctaccgtctggccctacacgtctaccagcatggcctgactggcttcctaggccaggt 161
          | |||||||||||||||||| |||||||| |||||||||||||||| ||||||||
Qy    945 CAGCTACCGTCTGGCCCTACACATCTACCAGCGTGGCCTGACTGGCTTCCTGGGCCAGGT 1004

Db    162 gacccgcttcgtggtcgacttcatgctgcatcactgcattgcccggtggattgcacagag 221
          ||||||||| ||||| ||||||||||||||||||||||||||||||||||||||||||
Qy   1005 GACCCGCTTTGTGGTGGACTTCATGCTGCATCACTGCATTGCCCGGTGGATTGCACAGAG 1064

Db    222 gggtggctgggtg 234
          |||||||||||||
Qy   1065 GGGTGGCTGGGTG 1077


RESULT    8
LOCUS       D88396S1      140 bp    DNA             PRI       05-DEC-1996
DEFINITION  Human DNA for apotosis regulator Bak, exon 1.
ACCESSION   D88396
NID         g1655491
KEYWORDS    apotosis-regulator; Bak.
SEGMENT     1 of 2
SOURCE      Homo sapiens DNA, clone:pGEMBak01-02.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 140)
  AUTHORS   Eguchi,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-OCT-1996) to the DDBJ/EMBL/GenBank databases.
            Hidetaka Eguchi, Saitama Cancer Center Research Institute,
            Depertment of Biochemistry; 818 Komuro, Ina, Kita-adachi-gun,
            Saitama 362, Japan (E-mail:hide@saitama-cc.go.jp,
            Tel:048-722-1111(ex.255), Fax:048-722-1739)
REFERENCE   2  (bases 1 to 140)
  AUTHORS   Eguchi,H. and Hayashi,S.
  TITLE     Estrogen alters expression of apoptosis-regulators, Bcl-2, Bcl-xL
            and Bak, as well as susceptibility to therapeutic agents of human
            breast cancer cells
  JOURNAL   Unpublished (1996)
FEATURES             Location/Qualifiers
     source          1..140
                     /organism="Homo sapiens"
                     /clone="pGEMBak01-02"
     exon            1..90
                     /number=1
     intron          91..140
                     /number=1
BASE COUNT       33 a     52 c     27 g     28 t
ORIGIN

  Query Match           6.8%;  Score 88;  DB 70;  Length 140;
  Best Local Similarity  98.9%;  Pred. No. 4.96e-49;
  Matches    89;  Conservative    0;  Mismatches   1;  Indels   0;  Gaps   0;

Db      1 acgctatgactcagagttccagaccatgttgcagcacctgcagcccacggcagagaatgc 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    804 ACGCTATGACTCAGAGTTCCAGACCATGTTGCAGCACCTGCAGCCCACGGCAGAGAATGC 863

Db     61 ctatgagtacttcaccaagattgccaccag 90
          |||||||||||||||||||||||||| ||||
Qy    864 CTATGAGTACTTCACCAAGATTGCCTCCAG 893


RESULT    9
LOCUS       I28278        215 bp    DNA             PAT       30-OCT-1996
DEFINITION  Sequence 5 from patent US 5569830.
ACCESSION   I28278
```

```
NID         g1819054
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 215)
  AUTHORS   Bennett,A., Labavitch,J.M., Powell,A. and Stotz,H.
  TITLE     Plant inhibitors of fungal polygalacturonases and their use to
            control fungal disease
  JOURNAL   Patent: US 5569830-A 5 29-OCT-1996;

FEATURES             Location/Qualifiers
     source          1..215
                     /organism="unknown"
BASE COUNT       15 a      8 c     25 g     26 t    141 others
ORIGIN

  Query Match            2.9%;  Score 37;  DB 57;  Length 215;
  Best Local Similarity  16.0%;  Pred. No. 4.92e-08;
  Matches    27;  Conservative   71;  Mismatches  69;  Indels   2;  Gaps   2;

Db      4 vtmssssvvsrtascndkakkdgnttsswttdccnrtwgvcdtdttyrvnndsghnkyss 63
          :|:  ::::::| :| ::| ::| ||:::  :|| : :|:|  : |: :  : |   : :
Qy    967 ATCTACCAGCGTGGCCTGACTGGCTTCCTGGGCCAGGTGACCCGCTTTGTGGTGGACTTC 1026

Db     64 anynyggnnvgaakthyythtnvsgadsktvtdsynasgtsssnggtdgnrsgadsygss 123
          | :|   :  : ::: :  ::  : :|: ::  |:| ::  | |:|  :| ::: :
Qy   1027 ATGCTGCATCACTGCATTGCCCGGTGGATTGCACAGAGG-GGTGGCTGGGTGGCAGCCCT 1085

Db    124 ktamtsrnrtgktannavdsrnmgdasvgsdkntkkhaknsadgkvgsk 172
          : |:| : :   |    : :: : :  : ::: |::: :   :|:: ::
Qy   1086 GAACTTGGGCAATGGTCC-CATCCTGAACGTGCTGGTGGTTCTGGGTGT 1133


RESULT   10
LOCUS       I28278        215 bp    DNA             PAT       30-OCT-1996
DEFINITION  Sequence 5 from patent US 5569830.
ACCESSION   I28278
NID         g1819054
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 215)
  AUTHORS   Bennett,A., Labavitch,J.M., Powell,A. and Stotz,H.
  TITLE     Plant inhibitors of fungal polygalacturonases and their use to
            control fungal disease
  JOURNAL   Patent: US 5569830-A 5 29-OCT-1996;

FEATURES             Location/Qualifiers
     source          1..215
                     /organism="unknown"
BASE COUNT       15 a      8 c     25 g     26 t    141 others
ORIGIN

  Query Match            2.7%;  Score 35;  DB 57;  Length 215;
  Best Local Similarity  11.2%;  Pred. No. 1.16e-06;
  Matches    23;  Conservative   85;  Mismatches  96;  Indels   1;  Gaps   1;

Db      1 mtnvtmssssvvsrtascndkakkdgnttsswttdccnrtwgvcdtdttyrvnndsghnk 60
          :| :|:: : :: : |:  ::  ::|   ::: | |  : : : :  |  ::  :: : :
Cp    621 CTGCTCCTCAGAAGCAGAGGGCAGGGCAGGCTCTCCGCACTCCTGCCTGGGAGGACCTGG 562

Db     61 yssan-ynyggnnvgaakthyythtnvsgadsktvtdsynasgtsssnggtdgnrsgads 119
           ::   : : |  :  |:|::: :   :  ::: : ::   :|  :    |:  ::||::
Cp    561 GCCTTGCCCCGAAGCCATTTTTCAGGTCTCAGTGGAGGACGGGATCAGCCTGCCGGGATC 502

Db    120 ygssktamtsrnrtgktannavdsrnmgdasvgsdkntkkhaknsadgkvgsknngdrnn 179
          : ::   : :  : |:     ::::  |: ::| ::  :::   :|:|::  :  |:
Cp    501 CTGGCCCAACCCGGGTGGCTCAGCAGGGTGGAGATGGAGGTCCCGAGGGGCTGGCGGCTG 442

Db    180 rygtgtksnvsnncgggnkrdvssy 204
           :|    : ::  | |  :::: ::
```

```
Cp    441 CTGCTCCCAGGGGCTGAGTGGGAGC 417


RESULT   11
LOCUS       HUMPECAM16    250 bp    DNA             PRI       07-JAN-1995
DEFINITION  Homo sapiens platelet/endothelial cell adhesion molecule-1
            (PECAM-1) gene, partial intron 13.
ACCESSION   L34646
NID         g598210
KEYWORDS    platelet endothelial cell adhesion molecule 1.
SEGMENT     16 of 27
SOURCE      Homo sapiens DNA.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 250)
  AUTHORS   Newman,P.J., Berndt,M.C., Gorski,J., White,G.C. II., Lyman,S.,
            Paddock,C. and Muller,W.A.
  TITLE     PECAM-1 (CD31) cloning and relation to adhesion molecules of the
            immunoglobulin gene superfamily
  JOURNAL   Science 247 (4947), 1219-1222 (1990)
  MEDLINE   90193682
REFERENCE   2  (bases 1 to 250)
  AUTHORS   Kirschbaum,N.E., Gumina,R.J. and Newman,P.J.
  TITLE     Organization of the gene for human platelet/endothelial cell
            adhesion molecule-1 shows alternatively spliced isoforms and a
            functionally complex cytoplasmic domain
  JOURNAL   Blood 84, 4028-4037 (1994)
  MEDLINE   95086194
FEATURES             Location/Qualifiers
     source          1..250
                     /organism="Homo sapiens"
                     /cell_type="leukocyte"
     intron          <1..>250
                     /gene="PECAM-1"
                     /note="putative"
                     /citation=[2]
                     /number=13
BASE COUNT       77 a      15 c     21 g    137 t
ORIGIN

  Query Match            2.1%;  Score 27;  DB 82;  Length 250;
  Best Local Similarity  66.7%;  Pred. No. 1.44e-01;
  Matches    54;  Conservative    0;  Mismatches  27;  Indels   0;  Gaps   0;

Db    100 ttatttttttatttttatatatttatatataatatttttatatttatattatttatatat 159
          || |||||  |||| || | ||| |  | ||  |  |||  |||| |||||  ||||| |
Cp    104 TTGTTTTTACTTTTCATCTTTTTCAGTTCTATAACATTTGAATTTTTATTAAGTATATTT 45

Db    160 ttttattttatatatatata 180
            | |||| | | |||| |||
Cp     44 CATGTATTGTCTCTATAAATA 24


RESULT   12
LOCUS       TTG8SCRNA    1049 bp    DNA             INV       14-FEB-1992
DEFINITION  T.thermophila G8-scRNA DNA.
ACCESSION   X57037
NID         g10810
KEYWORDS    heat shock induced; RNA polymerase III; scRNA.
SOURCE      Tetrahymena thermophila.
  ORGANISM  Tetrahymena thermophila
            Eukaryotae; mitochondrial eukaryotes; Alveolata; Ciliophora;
            Oligohymenophorea; Hymenostomatida; Tetrahymenina; Tetrahymena.
REFERENCE   1  (bases 1 to 1049)
  AUTHORS   Hallberg,R.L.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-DEC-1990) R.L. Hallberg, SYRACUSE UNIVERSITY, BIOLOGY
            DEPARTMENT, 130 COLLEGE PLACE, SYRACUSE NY 13244, USA
REFERENCE   2  (bases 1 to 1049)
  AUTHORS   Hallberg,E.M., Fung,P. and Hallberg,R.L.
  TITLE     Genomic sequence encoding a heat shock-induced, RNA polymerase
            III-transcribed RNA from Tetrahymena thermophila
```

```
  JOURNAL   Nucleic Acids Res. 20 (4), 912 (1992)
  MEDLINE   92178994
COMMENT     See also M14555.
FEATURES             Location/Qualifiers
     source          1..1049
                     /organism="Tetrahymena thermophila"
                     /strain="BIV"
     misc_feature    299..325
                     /note="heat shock elements (2)"
     misc_feature    390..405
                     /note="proximal sequence element"
     misc_feature    424..432
                     /note="T/A rich element"
     scRNA           447..752
                     /note="G8-SCRNA"
BASE COUNT      472 a    112 c    116 g    349 t
ORIGIN

  Query Match           2.1%;  Score 27;  DB 45;  Length 1049;
  Best Local Similarity  70.0%;  Pred. No. 1.44e-01;
  Matches    56;  Conservative    0;  Mismatches  23;  Indels   1;  Gaps   1;

Db    441 attaaaataaaaaaaaaaataaataaacttaataaaaaaacaactgctaaa-aactaaaa 499
          ||| | | | | || | |  ||||| |||||||||||||  ||| || || | |||| |||
Qy     25 ATTTATAGAGACAATACATGAAATATACTTAATAAAAATTCAAATGTTATAGAACTGAAA 84

Db    500 tgggtggaaagaaagatcaa 519
          | || |||| || | |||
Qy     85 AAGATGAAAAGTAAAAACAA 104


RESULT   13
LOCUS       PFTUBAI      2648 bp    DNA             INV       25-MAR-1993
DEFINITION  Plasmodium falciparum alpha-tubulin I gene.
ACCESSION   X15979
NID         g9979
KEYWORDS    alpha-tubulin; tubulin.
SOURCE      malaria parasite.
  ORGANISM  Plasmodium falciparum
            Eukaryotae; mitochondrial eukaryotes; Alveolata; Apicomplexa;
            Haemosporida; Plasmodium.
REFERENCE   1  (bases 1 to 2648)
  AUTHORS   Hyde,J.E.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-JUL-1989) Hyde J.E., University of Manchester
            Institute of Science and Technology, Dept of Biochemistry and
            Applied Molecular Biology, UMIST P O Box 88, Manchester M60 1QD, UK
REFERENCE   2  (bases 1 to 2648)
  AUTHORS   Holloway,S.P., Sims,P.F., Delves,C.J., Scaife,J.G. and Hyde,J.E.
  TITLE     Isolation of alpha-tubulin genes from the human malaria parasite,
            Plasmodium falciparum: sequence analysis of alpha-tubulin
  JOURNAL   Mol. Microbiol. 3 (11), 1501-1510 (1989)
  MEDLINE   90136080
FEATURES             Location/Qualifiers
     source          1..2648
                     /organism="Plasmodium falciparum"
                     /strain="K1 (Thailand)"
                     /cell_type="intraerythrocytic"
                     /clone_lib="lambda NM1149, HindIII"
                     /clone="H3.7"
     exon            <626..687
                     /number=1
     CDS             join(626..687,1042..1162,1347..2525)
                     /codon_start=1
                     /product="alpha-tubulin"
                     /db_xref="PID:g9980"
                     /db_xref="SWISS-PROT:P14642"
                     /translation="MREVISIHVGQAGIQVGNACWELFCLEHGIQPDGQMPSDKASRA
                     NDDAFNTFFSETGAGKHVPRCVFVDLEPTVVDEVRTGTYRQLFHPEQLISGKEDAANN
                     FARGHYTIGKEVIDVCLDRIRKLADNCTGLQGFLMFSAVGGGTGSGFGCLMLERLSVD
                     YGKKSKLNFCCWPSPQVSTAVVEPYNSVLSTHSLLEHTDVAIMLDNEAIYDICRRNLD
                     IERPTYTNLNRLIAQVISSLTASLRFDGALNVDVTEFQTNLVPYPRIHFMLSSYAPVV
                     SAEKAYHEQLSVSEITNSAFEPANMMAKCDPRHGKYMACCLMYRGDVVPKDVNAAVAT
                     IKTKRTIQFVDWCPTGFKCGINYQPPTVVPGGDLAKVMRAVCMISNSTAIAEVFSRMD
                     QKFDLMYAKRAFVHWYVGEGMEEGEFSEAREDLAALEKDYEEVGIESNEAEGEDEGYE
                     ADY"
     intron          688..1041
                     /number=1
     exon            1042..1162
                     /number=2
     intron          1163..1346
                     /number=2
     exon            1347..>2525
                     /number=3
BASE COUNT     1019 a    271 c    368 g    990 t
ORIGIN

  Query Match           2.1%;  Score 27;  DB 44;  Length 2648;
  Best Local Similarity  67.3%;  Pred. No. 1.44e-01;
  Matches    72;  Conservative    0;  Mismatches  33;  Indels   2;  Gaps   2;

Db    305 taatatatataaaagtataatttttttacaaataattaat-agatttatttttataagaat 363
          ||||||| || || ||    | ||| || | | |||  || | || || || |||| |||
Qy      4 TAATATAAATTAATGTGCTCTATTTATAGAGACAATACATGAAATATACTTAATAAAAAT 63

Db    364 taaaaa-taataaaattaaaaaaaaaaaaaaagaaaaatttacatatt 409
          | |||  | ||| || | ||||| |  ||||| ||||   || ||||
Qy     64 TCAAATGTTATAGAACTGAAAAAGATGAAAAGTAAAAACAACCTATT 110


RESULT   14
LOCUS       XLPOMCA      9747 bp    DNA             VRT       02-DEC-1993
DEFINITION  X.laevis POMC-A gene for proopiomelanocortin-A.
ACCESSION   X59370 S35811
NID         g64988
KEYWORDS    colour adaptation; corticotrope (ATCH); melanotrope (a-MSH);
            neuropeptide/opioid; opioid (beta-endorphin); POMC-A gene;
            preprohormone; proopiomelanocortin-A; repetitive element JH12;
            repetitive element Vi; secretion.
SOURCE      African clawed frog.
  ORGANISM  Xenopus laevis
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea;
            Pipidae; Xenopodinae; Xenopus.
REFERENCE   1  (bases 1 to 9747)
  AUTHORS   Deen,P.M.T.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-MAY-1991) P.M.T. Deen, Univ of Nijmegen,
            Toernnooiveld, 6525 ED Nijmegen, THE NETHERLANDS
REFERENCE   2  (bases 1 to 9747)
  AUTHORS   Deen,P.M.T., Terwel,D., Bussemakers,M.J.M., Roubos,E.W. and
            Martens,G.J.M.
  TITLE     Comparative analysis of the transcriptionally active
            Proopiomelanocortin genes A and B of Xenopus laevis
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 9747)
  AUTHORS   Deen,P.M., Roubos,E.W. and Martens,G.J.
  TITLE     Presence of Vi-transposon-like elements in the proopiomelanocortin
            gene A of Xenopus laevis does not affect gene activity
  JOURNAL   Mol. Gen. Genet. 230 (3), 491-493 (1991)
  MEDLINE   92114880
REFERENCE   4  (bases 1 to 9747)
  AUTHORS   Deen,P.M., Bussemakers,M.J., Terwel,D., Roubos,E.W. and
            Martens,G.J.
  TITLE     Comparative structural analysis of the transcriptionally active
            proopiomelanocortin genes A and B of Xenopus laevis
  JOURNAL   Mol. Biol. Evol. 9 (3), 483-494 (1992)
  MEDLINE   92261311
FEATURES             Location/Qualifiers
     source          1..9747
                     /organism="Xenopus laevis"
                     /clone_lib="lambda EMBL4"
                     /clone="lambda XPA5"
     promoter        1..490
     TATA_signal     462..469
     exon            491..538
```

```
                    /gene="POMC A"
                    /number=1
     CDS            join(491..538,4906..5048,7911..8922)
                    /gene="POMC A"
                    /product="proopiomelanocortin-A"
     intron         539..4905
                    /gene="POMC A"
                    /number=1
     misc_feature   1724..1957
                    /gene="POMC A"
                    /note="simple sequence"
     repeat_region  1972..2459
                    /note="Vi repetitive element"
     repeat_region  2471..2796
                    /note="JH12 repetitive element"
     repeat_region  3105..3504
                    /note="Vi repetitive element"
     repeat_region  4306..4590
                    /note="Vi repetitive element"
     repeat_region  4590..4841
                    /note="JH12 repetitive element"
     exon           4906..5048
                    /gene="POMC A"
                    /number=2
     intron         5049..7910
                    /gene="POMC A"
                    /number=2
     repeat_region  6886..7347
                    /note="Vi repetitive element"
     exon           7911..8922
                    /gene="POMC A"
                    /number=3
     repeat_region  9027..9144
                    /note="internal repeat"
     misc_feature   9438..9747
                    /note="simple sequence"
BASE COUNT     3228 a   1558 c   1715 g   3246 t
ORIGIN

  Query Match             2.1%;  Score 27;  DB 52;  Length 9747;
  Best Local Similarity  74.2%;  Pred. No. 1.44e-01;
  Matches    49;  Conservative    0;  Mismatches  16;  Indels   1;  Gaps   1;

Db   1562 tttttactttcattctttttcatatctataa-atgtatattttatttattcatatttatt 1620
          ||||||||||   |||||||||  ||||||| || |  |||||  |||   ||||||  |
Cp    101 TTTTTACTTTTCATCTTTTTCAGTTCTATAACATTTGAATTTTTATTAAGTATATTTCAT 42

Db   1621 gttttg 1626
          || |||
Cp     41 GTATTG 36


RESULT   15
LOCUS       XPU65216      610 bp    DNA             INV       03-JAN-1997
DEFINITION  Xenos pecki 28S ribosomal RNA gene, partial sequence.
ACCESSION   U65216
NID         g1762761
KEYWORDS    .
SOURCE      Xenos pecki.
  ORGANISM  Xenos pecki
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Arthropoda;
            Tracheata; Hexapoda; Insecta; Pterygota; Strepsiptera; Stylopoidea;
            Stylopidae; Xeninae; Xenos.
REFERENCE   1  (bases 1 to 610)
  AUTHORS   Whiting,M.F., Carpenter,J.C., Wheeler,Q.D. and Wheeler,W.C.
  TITLE     The Strepsiptera Problem: Phylogeny of the Holometabolous Insect
            Orders Inferred From 18S and 28S Ribosomal DNA Sequences and
            Morphology
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 610)
  AUTHORS   Whiting,M.F., Carpenter,J.C., Wheeler,Q.D. and Wheeler,W.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-JUL-1996) Entomology, American Museum of Natural
            History, Central Park West at 79th Street, New York, NY 10024, USA
FEATURES             Location/Qualifiers
     source          1..610
                     /organism="Xenos pecki"
     rRNA            <1..>610
                     /product="28S ribosomal RNA"
BASE COUNT      225 a     68 c     97 g    220 t
ORIGIN

  Query Match             2.0%;  Score 26;  DB 45;  Length 610;
  Best Local Similarity  70.3%;  Pred. No. 5.48e-01;
  Matches    45;  Conservative    0;  Mismatches  19;  Indels   0;  Gaps   0;

Db     67 tttttacttttctttttttttgtttttaaatttaaatagaaataaaatttttgctt 126
          |||||||||||| | |||||  ||| | |||  ||| |||    || || | | || | |
Cp    101 TTTTTACTTTTCATCTTTTTCAGTTCTATAACATTTGAATTTTTATTAAGTATATTTCAT 42

Db    127 gtat 130
          ||||
Cp     41 GTAT 38


Search completed: Thu Aug 21 13:16:30 1997
Job time : 2759 secs.
```

```
******************************************************************
MPSRCH (TM)
******************************************************************
```

Release 2.1D John F. Collins, Biocomputing Research Unit.


```
MPsrch_nn   n.a. - n.a. database search, using Smith-Waterman algorithm

Run on:          Thu Aug 21 13:16:51 1997;  MasPar time 135.57 Seconds
                                            934.038 Million cell updates/sec
Tabular output not generated.

Title:           >US-08-320-157-8
Description:     (1-1287) from US08320157.seq
Perfect Score:   1287
N.A. Sequence:      1 TTTTAATATAAATTAATGTG..........CCTCAAGAGTACAGAAGCTT 1287
     Comp:            AAAATTATATTTAATTACAC..........GGAGTTCTCATGTCTTCGAA

Scoring table:   TABLE default
                 Gap 6

Nmatch   STD :   Dbase 0;  Query 0

Searched:        134151 seqs, 49196315 bases x 2

Post-processing: Minimum Match 0%
                 Listing first 45 summaries

Database:        n-geneseq27
                   1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
                   8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
                   14:part14 15:part15 16:part16 17:part17 18:part18
                   19:part19 20:part20 21:part21 22:part22 23:part23
                   24:part24 25:part25 26:part26 27:part27

Statistics:      Mean 9.436;  Variance 5.783;  scale 1.632
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description	Pred. No.
	1	1280	99.5	1286	24	T42139	Bak-2 gene.	0.00e+00
	2	1269	98.6	6511	14	Q95493	Human Cdn-2 DNA.	0.00e+00
	3	890	69.2	2094	24	T42138	Bak gene.	0.00e+00
	4	889	69.1	2072	14	Q95492	Human Cdn-1 cDNA.	0.00e+00
	5	768	59.7	1968	19	T17375	Bcl-Y cDNA.	0.00e+00
	6	753	58.5	5408	14	Q95494	Human Cdn-3 DNA.	0.00e+00
c	7	88	6.8	1047	2	Q10572	Human Natriuretic Pep	5.85e-38
	8	77	6.0	1047	2	Q10572	Human Natriuretic Pep	5.42e-31
c	9	45	3.5	204	1	N81164	Base substituted E.co	1.24e-11
	10	43	3.3	91	9	Q51746	Oligonucleotide probe	1.70e-10
c	11	42	3.3	91	9	Q51746	Oligonucleotide probe	6.22e-10
	12	40	3.1	204	1	N81164	Base substituted E.co	8.16e-09
	13	36	2.8	114	12	Q70468	Generic DNA sequence	1.26e-06
	14	36	2.8	114	12	Q70466	Generic DNA sequence	1.26e-06
	15	36	2.8	114	12	Q70467	Generic DNA sequence	1.26e-06
c	16	35	2.7	114	12	Q70465	Generic DNA sequence	4.33e-06
	17	34	2.6	114	12	Q70465	Generic DNA sequence	1.47e-05
	18	33	2.6	114	12	Q70469	Generic DNA sequence	4.95e-05
c	19	34	2.6	114	12	Q70467	Generic DNA sequence	1.47e-05
c	20	34	2.6	114	12	Q70468	Generic DNA sequence	1.47e-05
c	21	33	2.6	114	12	Q70466	Generic DNA sequence	4.95e-05
	22	31	2.4	114	12	Q70470	Generic DNA sequence	5.38e-04
	23	31	2.4	114	12	Q70472	Generic DNA sequence	5.38e-04
c	24	31	2.4	114	12	Q70470	Generic DNA sequence	5.38e-04
c	25	31	2.4	114	12	Q70469	Generic DNA sequence	5.38e-04
c	26	30	2.3	565	6	Q35072	HCV envelope region n	1.74e-03
	27	28	2.2	70	24	T14325	Conjugate formed by l	1.73e-02
	28	28	2.2	114	12	Q70471	Generic DNA sequence	1.73e-02
	29	28	2.2	114	12	Q70473	Generic DNA sequence	1.73e-02
c	30	28	2.2	114	12	Q70471	Generic DNA sequence	1.73e-02
c	31	28	2.2	114	12	Q70472	Generic DNA sequence	1.73e-02
	32	27	2.1	565	6	Q35072	HCV envelope region n	5.32e-02
c	33	26	2.0	36	2	Q11195	Ballast Constituent c	1.61e-01
c	34	26	2.0	39	7	Q51787	Mixed oligonucleotide	1.61e-01
c	35	26	2.0	114	12	Q70473	Generic DNA sequence	1.61e-01
	36	26	2.0	3871	2	N71302	HSV-1 gB and surround	1.61e-01
c	37	24	1.9	39	24	T14323	Sequence used in the	1.38e+00
	38	25	1.9	75	21	T13612	DC43 TSAR library gen	4.76e-01
	39	25	1.9	82	21	T13610	DC43 TSAR library gen	4.76e-01
	40	25	1.9	198	7	Q42784	Ligand-induced gene,	4.76e-01
	41	24	1.9	501	3	N50026	Sequence encoding new	1.38e+00
	42	24	1.9	501	3	N50030	Sequence encoding new	1.38e+00
	43	24	1.9	501	3	N50024	Sequence encoding new	1.38e+00
	44	24	1.9	2456	9	Q51137	Retinoblastoma-associ	1.38e+00
	45	24	1.9	2492	10	Q67603	Retinoblastoma protei	1.38e+00

ALIGNMENTS

```
RESULT    1
ID   T42139 standard; cDNA; 1286 BP.
AC   T42139;
DT   22-FEB-1997  (first entry)
DE   Bak-2 gene.
KW   Human; Bak-2; apoptosis; latency; virus replication;
KW   Epstein-Barr virus; BHRF1; fusion protein; epitope tag;
KW   drug screening; co-precipitation; ELISA; immunoassay; antibody;
KW   protein interactive trapping; virucide; antitumour; diagnostic; ds.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   5'UTR           1..543
FT   /*tag= a
FT   CDS             544..1179
FT   /*tag= b
FT   /product= Bak-2 protein
FT   3'UTR           1180..1286
FT   /*tag= c
PN   WO9633416-A1.
PD   24-OCT-1996.
PF   19-APR-1996; U05639.
PR   20-APR-1995; US-426529.
PA   (LXRB-) LXR BIOTECHNOLOGY INC.
PI   Barr PJ,  Kiefer MC;
DR   WPI; 96-485886/48.
DR   P-PSDB; W03669.
PT   Screening for anti-viral agents - by detecting the ability of an
PT   agent to disrupt the interaction of a Bak protein and a viral
PT   protein
PS   Disclosure; Fig 2; 24pp; English.
CC   The sequence encodes Bak-2 protein, which is a bcl-1 homologue which
CC   interacts with Epstein-Barr virus (EBV) early lytic cycle BHRF1
CC   protein, and is capable of modulating apoptosis.  The Bak-2 gene is
CC   located on human chromosome-20.  The protein may be used in
CC   complete or partial form, or as an epitope tag fusion protein, in a
CC   new virucide drug screening method, which involves combination of
CC   Bak-2 protein and a viral protein (e.g. EBV BHRF1), exposure to a
CC   test compound, and monitoring for disruption of the interaction,
CC   e.g. by co-precipitation, protein interactive trapping or ELISA.
CC   Interaction of Bak-2 and viral proteins allows viral replication or
```

```
CC   latency in the absence of apoptosis.  Compounds which inhibit the
CC   interaction may be used as virucide, antitumour or diagnostic agents.
SQ   Sequence   1286 BP;    298 A;    341 C;    328 G;    319 T;

  Query Match           99.5%;  Score 1280;  DB 24;  Length 1286;
  Best Local Similarity  99.9%;  Pred. No. 0.00e+00;
  Matches  1286;  Conservative    0;  Mismatches   0;  Indels   1;  Gaps   1;

Db      1 ttttaatataaattaatgtgctctatttatagagacaatacatgaaatatacttaataaa 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      1 TTTTAATATAAATTAATGTGCTCTATTTATAGAGACAATACATGAAATATACTTAATAAA 60

Db     61 aattcaaatgttatagaactgaaaaagatgaaaagtaaaaacaacctattccccagaggt 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     61 AATTCAAATGTTATAGAACTGAAAAAGATGAAAAGTAAAAACAACCTATTCCCCAGAGGT 120

Db    121 agccactgtccatagtttctattttagattctttcctttatacaagattattatagcttc 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    121 AGCCACTGTCCATAGTTTCTATTTTAGATTCTTTCCTTTATACAAGATTATTATAGCTTC 180

Db    181 tatttttggtgtatgaactgtagtcctagaggattttattagttatgagttctataact 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    181 TATTTTTGGTGTATGAACTGTAGTCCTAGAGGATTTTATTAGTTATGAGTTCTATAACT 240

Db    241 aagatccatcatcttagttgctaagaacgtagatactgagaacatcatttaaaaaaacat 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    241 AAGATCCATCATCTTAGTTGCTAAGAACGTAGATACTGAGAACATCATTTAAAAAAACAT 300

Db    301 ttttggctggcacctcatgatcactggagtctcgcgggtccctcaggctgcacagggaca 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    301 TTTTGGCTGGCACCTCATGATCACTGGAGTCTCGCGGGTCCCTCAGGCTGCACAGGGACA 360

Db    361 agtaaaggctacatccagatgctgggaatgcactgacgcccattcctggaaactgggctc 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    361 AGTAAAGGCTACATCCAGATGCTGGGAATGCACTGACGCCCATTCCTGGAAACTGGGCTC 420

Db    421 ccactcagcccctgggagcagcagccgccagccccctcgggacctccatctccaccctgct 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    421 CCACTCAGCCCCTGGGAGCAGCAGCCGCCAGCCCCTCGGGACCTCCATCTCCACCCTGCT 480

Db    481 gagccacccgggttgggccaggatcccggcaggctgatcccgtcctccactgagacctga 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    481 GAGCCACCCGGGTTGGGCCAGGATCCCGGCAGGCTGATCCCGTCCTCCACTGAGACCTGA 540

Db    541 aaaatggcttcggggcaaggcccaggtcctcccaggcaggagtgcggagagcctgccctg 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    541 AAAATGGCTTCGGGGCAAGGCCCAGGTCCTCCCAGGCAGGAGTGCGGAGAGCCTGCCCTG 600

Db    601 ccctctgcttctgaggagcaggtagcccaggacacagaggaggttttccgcagctacgtt 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    601 CCCTCTGCTTCTGAGGAGCAGGTAGCCCAGGACACAGAGGAGGTTTTCCGCAGCTACGTT 660

Db    661 ttttaccaccatcagcaggaacaggaggctgaaggggcggctgcccctgccgacccagag 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    661 TTTTACCACCATCAGCAGGAACAGGAGGCTGAAGGGGCGGCTGCCCCTGCCGACCCAGAG 720

Db    721 atggtcaccttacctctgcaacctagcagcaccatggggcaggtgggacggcagctcgcc 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    721 ATGGTCACCTTACCTCTGCAACCTAGCAGCACCATGGGGCAGGTGGGACGGCAGCTCGCC 780

Db    781 atcattggggacgacatcaaccgacgctatgactcagagttccagaccatgttgcagcac 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    781 ATCATTGGGGACGACATCAACCGACGCTATGACTCAGAGTTCCAGACCATGTTGCAGCAC 840

Db    841 ctgcagcccacggcagagaatgcctatgagtacttcaccaagattgcctccagcctgttt 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    841 CTGCAGCCCACGGCAGAGAATGCCTATGAGTACTTCACCAAGATTGCCTCCAGCCTGTTT 900

Db    901 gagagtggcatcaattggggccgtgtggtggctcttctgggcttcagctaccgtctggcc 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    901 GAGAGTGGCATCAATTGGGGCCGTGTGGTGGCTCTTCTGGGCTTCAGCTACCGTCTGGCC 960

Db    961 ctacacatctaccagcgtggcctgactggcttcctgggccaggtgacccgctttgtggtg 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    961 CTACACATCTACCAGCGTGGCCTGACTGGCTTCCTGGGCCAGGTGACCCGCTTTGTGGTG 1020

Db   1021 gacttcatgctgcatcactgcattgcccggtggattgcacagaggggtggctgggtggca 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1021 GACTTCATGCTGCATCACTGCATTGCCCGGTGGATTGCACAGAGGGGTGGCTGGGTGGCA 1080

Db   1081 gccctgaacttgggcaatggtcccatcctgaacgtgctggtggttctgggtgtggttctg 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1081 GCCCTGAACTTGGGCAATGGTCCCATCCTGAACGTGCTGGTGGTTCTGGGTGTGGTTCTG 1140

Db   1141 ttgggccagtttgtggtacgaagattcttcaaatcatgactcccaagggtgcctttgggg 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1141 TTGGGCCAGTTTGTGGTACGAAGATTCTTCAAATCATGACTCCCAAGGGTGCCTTTGGGG 1200

Db   1201 tcccagttcagacccctgcctggacttaagcgaagtctttgccttctctgctccttgcag 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1201 TCCCAGTTCAGACCCCTGCCTGGACTTAAGCGAAGTCTTTGCCTTCTCTGCTCCTTGCAG 1260

Db   1261 g-tccccccctcaagagtacagaagctt 1286
          | |||||||||||||||||||||||||
Qy   1261 GGTCCCCCCCTCAAGAGTACAGAAGCTT 1287


RESULT    2
ID   Q95493 standard; DNA; 6511 BP.
AC   Q95493;
DT   21-NOV-1995  (first entry)
DE   Human Cdn-2 DNA.
KW   Cdn-2; apoptosis modulator; adoptive immunotherapy; therapy; HIV;
KW   autoimmune disease; reperfusion injury; hepatitis, osteoporosis;
KW   shock; lymphoma; eczema; ss.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   CDS             3312..3947
FT   /*tag=  a
PN   WO9515084-A.
PD   08-JUN-1995.
PF   30-NOV-1994; U13930.
PR   30-NOV-1993; US-160067.
PR   07-OCT-1994; US-320157.
PA   (LXRB-) LXR BIOTECHNOLOGY INC.
PI   Barr PJ,  Kiefer MC;
DR   WPI; 95-215106/28.
DR   P-PSDB; R77877.
PT   New nucleic acid sequences encoding Cdn apoptosis modulators  - and
PT   related vectors, transformed cells, proteins and antibodies, useful
PT   or diagnosis and treatment e.g. of HIV infection, reperfusion injury
PT   etc.
PS   Claim 6; Fig.5A-H; 66pp; English.
CC   Cdn-2 cDNA was isolated from a human placental genomic library
CC   using a 950 bp fragment of Cdn-1 cDNA.  Expression of Cdn-2
CC   in mouse progenitor B-cell FL5.12 cells decreased IL-3-induced
CC   apoptosis.  The Cdn-2 protein displayed 97% amino acid identity
CC   with Cdn-1 (R77876).
SQ   Sequence   6511 BP;    1513 A;    1620 C;    1605 G;    1773 T;

  Query Match           98.6%;  Score 1269;  DB 14;  Length 6511;
  Best Local Similarity  99.8%;  Pred. No. 0.00e+00;
  Matches  1287;  Conservative    0;  Mismatches   0;  Indels   3;  Gaps   3;

Db   2768 ttttaatataaattaatgtgctctatttatagagacaatacatgaaatatacttaataaa 2827
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      1 TTTTAATATAAATTAATGTGCTCTATTTATAGAGACAATACATGAAATATACTTAATAAA 60

Db   2828 aattcaaatgttatagaactgaaaaagatgaaaagtaaaaacaacctattccccagaggt 2887
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     61 AATTCAAATGTTATAGAACTGAAAAAGATGAAAAGTAAAAACAACCTATTCCCCAGAGGT 120

Db   2888 agccactgtccatagtttctattttagattctttcctttatacaagattattatagcttc 2947
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    121 AGCCACTGTCCATAGTTTCTATTTTAGATTCTTTCCTTTATACAAGATTATTATAGCTTC 180
```

```
Db   2948 tatttttggtgtatgaactgtagtcctagaggattttattagttatgagttctataact 3007
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    181 TATTTTTTGGTGTATGAACTGTAGTCCTAGAGGATTTTATTAGTTATGAGTTCTATAACT 240

Db   3008 aagatccatcatcttagttgctaagaacgtagatactgagaacatcatttaaaaaaacat 3067
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    241 AAGATCCATCATCTTAGTTGCTAAGAACGTAGATACTGAGAACATCATTTAAAAAAACAT 300

Db   3068 ttttggctggcacctctatgatcactggagtctcgcgggtccctcaggctgcacagggac 3127
          |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Qy    301 TTTTGGCTGGCACCTC-ATGATCACTGGAGTCTCGCGGGTCCCTCAGGCTGCACAGGGAC 359

Db   3128 aagtaaaggctacatccagatgctgggaatgcactgacgcccattcctggaaactgggct 3187
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    360 AAGTAAAGGCTACATCCAGATGCTGGGAATGCACTGACGCCCATTCCTGGAAACTGGGCT 419

Db   3188 cccactcagcccctgggagcagcagccgccagcccctcgggacctccatctccaccctgc 3247
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    420 CCCACTCAGCCCCTGGGAGCAGCAGCCGCCAGCCCCTCGGGACCTCCATCTCCACCCTGC 479

Db   3248 tgagccacccgggttgggccaggatcccggcaggctgatcccgtcctccactgagacctg 3307
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    480 TGAGCCACCCGGGTTGGGCCAGGATCCCGGCAGGCTGATCCCGTCCTCCACTGAGACCTG 539

Db   3308 aaaaatggcttcggggcaaggcccaggtcctcccaggcaggagtgcggagagcctgccct 3367
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    540 AAAAATGGCTTCGGGGCAAGGCCCAGGTCCTCCCAGGCAGGAGTGCGGAGAGCCTGCCCT 599

Db   3368 gccctctgcttctgaggagcaggtagcccaggacacagaggaggttttccgcagctacgt 3427
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    600 GCCCTCTGCTTCTGAGGAGCAGGTAGCCCAGGACACAGAGGAGGTTTTCCGCAGCTACGT 659

Db   3428 tttttaccaccatcagcaggaacaggaggctgaaggggcggctgcccctgccgacccaga 3487
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    660 TTTTTACCACCATCAGCAGGAACAGGAGGCTGAAGGGGCGGCTGCCCCTGCCGACCCAGA 719

Db   3488 gatggtcaccttacctctgcaacctagcagcaccatggggcaggtgggacggcagctcgc 3547
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    720 GATGGTCACCTTACCTCTGCAACCTAGCAGCACCATGGGGCAGGTGGGACGGCAGCTCGC 779

Db   3548 catcattggggacgacatcaaccgacgctatgactcagagttccagaccatgttgcagca 3607
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    780 CATCATTGGGGACGACATCAACCGACGCTATGACTCAGAGTTCCAGACCATGTTGCAGCA 839

Db   3608 cctgcagcccacggcagagaatgcctatgagtacttcaccaagattgcctccagcctgtt 3667
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    840 CCTGCAGCCCACGGCAGAGAATGCCTATGAGTACTTCACCAAGATTGCCTCCAGCCTGTT 899

Db   3668 tgagagtggcatcaattggggccgtgtggtggctcttctgggcttcagctaccgtctggc 3727
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    900 TGAGAGTGGCATCAATTGGGGCCGTGTGGTGGCTCTTCTGGGCTTCAGCTACCGTCTGGC 959

Db   3728 cctacacatctaccagcgtggcctgactggcttcctgggccaggtgacccgctttgtggt 3787
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    960 CCTACACATCTACCAGCGTGGCCTGACTGGCTTCCTGGGCCAGGTGACCCGCTTTGTGGT 1019

Db   3788 ggacttcatgctgcatcactgcattgcccggtggattgcacagaggggtggctgggtggc 3847
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    1020 GGACTTCATGCTGCATCACTGCATTGCCCGGTGGATTGCACAGAGGGGTGGCTGGGTGGC 1079

Db   3848 agccctgaacttgggcaatggtcccatcctgaacgtgctggtggttctgggtgtggttct 3907
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    1080 AGCCCTGAACTTGGGCAATGGTCCCATCCTGAACGTGCTGGTGGTTCTGGGTGTGGTTCT 1139

Db   3908 gttgggccagtttgtggtacgaagattcttcaaatcatgactcccaagggtgccctttgg 3967
          ||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Qy    1140 GTTGGGCCAGTTTGTGGTACGAAGATTCTTCAAATCATGACTCCCAAGGGTGCC-TTTGG 1198

Db   3968 ggtcccagttcagacccctgcctggacttaagcgaagtctttgccttctctgctccttgc 4027
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    1199 GGTCCCAGTTCAGACCCCTGCCTGGACTTAAGCGAAGTCTTTGCCTTCTCTGCTCCTTGC 1258

Db   4028 aggggtccccccctcaagagtacagaagctt 4057
```

```
          |||| |||||||||||||||||||||||||
Qy    1259 AGGG-TCCCCCCTCAAGAGTACAGAAGCTT 1287
```

RESULT 3
ID T42138 standard; cDNA; 2094 BP.
AC T42138;
DT 22-FEB-1997 (first entry)
DE Bak gene.
KW Human; Bak; apoptosis; latency; virus replication;
KW Epstein-Barr virus; BHRF1; fusion protein; epitope tag;
KW drug screening; co-precipitation; ELISA; immunoassay; antibody;
KW protein interactive trapping; virucide; antitumour; diagnostic; ss.
OS Homo sapiens.

FH	Key	Location/Qualifiers
FT	5'UTR	1..200
FT	/*tag= a	
FT	CDS	201..836
FT	/*tag= b	
FT	/product= Bak protein	
FT	3'UTR	837..2094
FT	/*tag= c	

PN WO9633416-A1.
PD 24-OCT-1996.
PF 19-APR-1996; U05639.
PR 20-APR-1995; US-426529.
PA (LXRB-) LXR BIOTECHNOLOGY INC.
PI Barr PJ, Kiefer MC;
DR WPI; 96-485886/48.
DR P-PSDB; W03668.
PT Screening for anti-viral agents - by detecting the ability of an
PT agent to disrupt the interaction of a Bak protein and a viral
PT protein
PS Disclosure; Fig 1; 24pp; English.
CC The sequence encodes Bak protein, which is a bcl-1 homologue which
CC interacts with Epstein-Barr virus (EBV) early lytic cycle BHRF1
CC protein, and is capable of modulating apoptosis. The Bak gene is
CC located on human chromosome-6 at 6p21-23. The protein may be used
CC in complete or partial form, or as an epitope tag fusion protein,
CC in a new virucide drug screening method, which involves combination
CC of Bak protein and a viral protein (e.g. EBV BHRF1), exposure to a
CC test compound, and monitoring for disruption of the interaction,
CC e.g. by co-precipitation, protein interactive trapping or ELISA.
CC Interaction of Bak and viral proteins allows viral replication or
CC latency in the absence of apoptosis. Compounds which inhibit the
CC interaction may be used as virucide, antitumour or diagnostic agents.
SQ Sequence 2094 BP; 410 A; 608 C; 606 G; 470 T;

Query Match 69.2%; Score 890; DB 24; Length 2094;
Best Local Similarity 98.2%; Pred. No. 0.00e+00;
Matches 917; Conservative 0; Mismatches 15; Indels 2; Gaps 2;

```
Db     13 gggacaagtaaaggctacatccagatgccgggaatgcactgacgcccattcctggaaact 72
          |||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Qy    355 GGGACAAGTAAAGGCTACATCCAGATGCTGGGAATGCACTGACGCCCATTCCTGGAAACT 414

Db     73 gggctcccactcagcccctgggagcagcagccgccagcccctcgg-acctccatctccac 131
          |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Qy    415 GGGCTCCCACTCAGCCCCTGGGAGCAGCAGCCGCCAGCCCCTCGGGACCTCCATCTCCAC 474

Db    132 cctgctgagccacccgggttgggccaggatcccggcaggctgatcccgtcctccactgag 191
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    475 CCTGCTGAGCCACCCGGGTTGGGCCAGGATCCCGGCAGGCTGATCCCGTCCTCCACTGAG 534

Db    192 acctgaaaaatggcttcggggcaaggcccaggtcctcccaggcaggagtgcggagagcct 251
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    535 ACCTGAAAAATGGCTTCGGGGCAAGGCCCAGGTCCTCCCAGGCAGGAGTGCGGAGAGCCT 594

Db    252 gccctgccctctgcttctgaggagcaggtagcccaggacacagaggaggttttccgcagc 311
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    595 GCCCTGCCCTCTGCTTCTGAGGAGCAGGTAGCCCAGGACACAGAGGAGGTTTTCCGCAGC 654

Db    312 tacgtttttttaccgccatcagcaggaacaggaggctgaaggggtggctgcccctgccgac 371
```

```
Qy    655 TACGTTTTTTACCACCATCAGCAGGAACAGGAGGCTGAAGGGGCGGCTGCCCCTGCCGAC 714

Db    372 ccagagatggtcaccttacctctgcaacctagcagcaccatggggcaggtgggacggcag 431
Qy    715 CCAGAGATGGTCACCTTACCTCTGCAACCTAGCAGCACCATGGGGCAGGTGGGACGGCAG 774

Db    432 ctcgccatcatcggggacgacatcaaccgacgctatgactcagagttccagaccatgttg 491
Qy    775 CTCGCCATCATTGGGGACGACATCAACCGACGCTATGACTCAGAGTTCCAGACCATGTTG 834

Db    492 cagcacctgcagcccacggcagagaatgcctatgagtacttcaccaagattgccaccagc 551
Qy    835 CAGCACCTGCAGCCCACGGCAGAGAATGCCTATGAGTACTTCACCAAGATTGCCTCCAGC 894

Db    552 ctgtttgagagtggcatcaattggggccgtgtggtggctcttctgggcttcggctaccgt 611
Qy    895 CTGTTTGAGAGTGGCATCAATTGGGGCCGTGTGGTGGCTCTTCTGGGCTTCAGCTACCGT 954

Db    612 ctggccctacacgtctaccagcatggcctgactggcttcctaggccaggtgacccgcttc 671
Qy    955 CTGGCCCTACACATCTACCAGCGTGGCCTGACTGGCTTCCTGGGCCAGGTGACCCGCTTT 1014

Db    672 gtggtcgacttcatgctgcatcactgcattgcccggtggattgcacagaggggtggctgg 731
Qy   1015 GTGGTGGACTTCATGCTGCATCACTGCATTGCCCGGTGGATTGCACAGAGGGGTGGCTGG 1074

Db    732 gtggcagccctgaacttgggcaatggtcccatcctgaacgtgctggtggttctgggtgtg 791
Qy   1075 GTGGCAGCCCTGAACTTGGGCAATGGTCCCATCCTGAACGTGCTGGTGGTTCTGGGTGTG 1134

Db    792 gttctgttgggccagtttgtggtacgaagattcttcaaatcatgactcccaagggtgccc 851
Qy   1135 GTTCTGTTGGGCCAGTTTGTGGTACGAAGATTCTTCAAATCATGACTCCCAAGGGTGCCT 1194

Db    852 tttgggtcccggttcagacccctgcctggacttaagcgaagtctttgccttctctgttcc 911
Qy   1195 TTGGGGTCCCAGTTCAGACCCCTGCCTGGACTTAAGCGAAGTCTTTGCCTTCTCTGCTCC 1254

Db    912 cttgcagggtccccccctcaagagtacagaagctt 945
Qy   1255 -TTGCAGGGTCCCCCCTCAAGAGTACAGAAGCTT 1287
```

```
RESULT    4
ID    Q95492 standard; cDNA; 2072 BP.
AC    Q95492;
DT    21-NOV-1995  (first entry)
DE    Human Cdn-1 cDNA.
KW    Cdn-1; apoptosis modulator; adoptive immunotherapy; therapy; HIV;
KW    autoimmune disease; reperfusion injury; hepatitis, osteoporosis;
KW    shock; lymphoma; eczema; ss.
OS    Homo sapiens.
FH    Key             Location/Qualifiers
FT    CDS             188..903
FT    /*tag= a
PN    WO9515084-A.
PD    08-JUN-1995.
PF    30-NOV-1994; U13930.
PR    30-NOV-1993; US-160067.
PR    07-OCT-1994; US-320157.
PA    (LXRB-) LXR BIOTECHNOLOGY INC.
PI    Barr PJ,  Kiefer MC;
DR    WPI; 95-215106/28.
DR    P-PSDB; R77876.
PT    New nucleic acid sequences encoding Cdn apoptosis modulators  - and
PT    related vectors, transformed cells, proteins and antibodies, useful
PT    or diagnosis and treatment e.g. of HIV infection, reperfusion injury
PT    etc.
PS    Claim 4; Fig.3A-C; 66pp; English.
CC    Cdn-1 cDNA was isolated from a human heart cDNA library using a
CC    previously isolated clone as probe.  Recombinant Cdn-1 was produced
CC    in Sf9 and human colon adenocarcinoma HT29 cells.  Expression of
CC    Cdn-1 in WI-L2 lymphoblastoid cells resulted in increased cell
CC    survival in response to anti-Fas-mediated apoptosis.
SQ    Sequence   2072 BP;    404 A;    603 C;    598 G;    467 T;

  Query Match           69.1%;  Score 889;  DB 14;  Length 2072;
  Best Local Similarity  98.2%;  Pred. No. 0.00e+00;
  Matches   916;  Conservative    0;  Mismatches  15;  Indels   2;  Gaps   2;

Db      1 ggacaagtaaaggctacatccagatgccgggaatgcactgacgcccattcctggaaactg 60
Qy    356 GGACAAGTAAAGGCTACATCCAGATGCTGGGAATGCACTGACGCCCATTCCTGGAAACTG 415

Db     61 ggctcccactcagcccctgggagcagcagccgccagcccctcgg-acctccatctccacc 119
Qy    416 GGCTCCCACTCAGCCCCTGGGAGCAGCAGCCGCCAGCCCCTCGGGACCTCCATCTCCACC 475

Db    120 ctgctgagccacccgggttgggccaggatcccggcaggctgatcccgtcctccactgaga 179
Qy    476 CTGCTGAGCCACCCGGGTTGGGCCAGGATCCCGGCAGGCTGATCCCGTCCTCCACTGAGA 535

Db    180 cctgaaaaatggcttcggggcaaggcccaggtcctcccaggcaggagtgcggagagcctg 239
Qy    536 CCTGAAAAATGGCTTCGGGGCAAGGCCCAGGTCCTCCCAGGCAGGAGTGCGGAGAGCCTG 595

Db    240 ccctgccctctgcttctgaggagcaggtagcccaggacacagaggaggttttccgcagct 299
Qy    596 CCCTGCCCTCTGCTTCTGAGGAGCAGGTAGCCCAGGACACAGAGGAGGTTTTCCGCAGCT 655

Db    300 acgtttttttaccgccatcagcaggaacaggaggctgaaggggtggctgcccctgccgacc 359
Qy    656 ACGTTTTTTACCACCATCAGCAGGAACAGGAGGCTGAAGGGGCGGCTGCCCCTGCCGACC 715

Db    360 cagagatggtcaccttacctctgcaacctagcagcaccatggggcaggtgggacggcagc 419
Qy    716 CAGAGATGGTCACCTTACCTCTGCAACCTAGCAGCACCATGGGGCAGGTGGGACGGCAGC 775

Db    420 tcgccatcatcggggacgacatcaaccgacgctatgactcagagttccagaccatgttgc 479
Qy    776 TCGCCATCATTGGGGACGACATCAACCGACGCTATGACTCAGAGTTCCAGACCATGTTGC 835

Db    480 agcacctgcagcccacggcagagaatgcctatgagtacttcaccaagattgccaccagcc 539
Qy    836 AGCACCTGCAGCCCACGGCAGAGAATGCCTATGAGTACTTCACCAAGATTGCCTCCAGCC 895

Db    540 tgtttgagagtggcatcaattggggccgtgtggtggctcttctgggcttcggctaccgtc 599
Qy    896 TGTTTGAGAGTGGCATCAATTGGGGCCGTGTGGTGGCTCTTCTGGGCTTCAGCTACCGTC 955

Db    600 tggccctacacgtctaccagcatggcctgactggcttcctaggccaggtgacccgcttcg 659
Qy    956 TGGCCCTACACATCTACCAGCGTGGCCTGACTGGCTTCCTGGGCCAGGTGACCCGCTTTG 1015

Db    660 tggtcgacttcatgctgcatcactgcattgcccggtggattgcacagaggggtggctggg 719
Qy   1016 TGGTGGACTTCATGCTGCATCACTGCATTGCCCGGTGGATTGCACAGAGGGGTGGCTGGG 1075

Db    720 tggcagccctgaacttgggcaatggtcccatcctgaacgtgctggtggttctgggtgtgg 779
Qy   1076 TGGCAGCCCTGAACTTGGGCAATGGTCCCATCCTGAACGTGCTGGTGGTTCTGGGTGTGG 1135

Db    780 ttctgttgggccagtttgtggtacgaagattcttcaaatcatgactcccaagggtgccct 839
Qy   1136 TTCTGTTGGGCCAGTTTGTGGTACGAAGATTCTTCAAATCATGACTCCCAAGGGTGCCTT 1195

Db    840 ttgggtcccggttcagacccctgcctggacttaagcgaagtctttgccttctctgttccc 899
Qy   1196 TGGGGTCCCAGTTCAGACCCCTGCCTGGACTTAAGCGAAGTCTTTGCCTTCTCTGCTCC- 1254

Db    900 ttgcagggtccccccctcaagagtacagaagctt 932
Qy   1255 TTGCAGGGTCCCCCCTCAAGAGTACAGAAGCTT 1287
```

```
RESULT    5
ID   T17375 standard; cDNA; 1968 BP.
AC   T17375;
DT   02-JUL-1996  (first entry)
DE   Bcl-Y cDNA.
KW   Bcl-Y; apoptosis; cell proliferation; cell death; diagnosis;
KW   gene therapy; ss.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   CDS             65..700
FT   /*tag=  a
PN   WO9605232-A1.
PD   22-FEB-1996.
PF   09-AUG-1995; U10103.
PR   09-AUG-1994; US-287427.
PR   11-OCT-1994; US-321071.
PA   (IMMU-) IMMUNOGEN INC.
PI   Chittenden TD;
DR   WPI; 96-139648/14.
DR   P-PSDB; R81451.
PT   New isolated human Bcl-Y protein - used to develop prods. for
PT   treating disorders characterised by inappropriate cell proliferation
PT   or cell death
PS   Claim 6; Fig 4; 100pp; English.
CC   A full-length cDNA clone (T17375) codes for Bcl-Y (R81451), a
CC   protein that induces apoptosis in cells and functions as a negative
CC   regulator of Bcl-2 function.  It was isolated from a Jurkat cell
CC   cDNA library using as probe a partial Bcl-Y cDNA clone obtd. by
CC   PCR of DNA derived from the Namalwa cell line.  The cDNA can be used
CC   for prodn. of recombinant Bcl-Y, as a probe, to produce transgenic
CC   animal models, and in the gene therapy of disorders characterised by
CC   inappropriate cell proliferation or cell death.
SQ   Sequence   1968 BP;    382 A;    560 C;    576 G;    450 T;

  Query Match           59.7%;  Score 768;  DB 19;  Length 1968;
  Best Local Similarity  98.0%;  Pred. No. 0.00e+00;
  Matches   794;  Conservative    0;  Mismatches  14;  Indels   2;  Gaps   2;

Db      1 tgagccacccgggttgggccaggatcccggcaggctgatcccgtcctccactgagacctg 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    480 TGAGCCACCCGGGTTGGGCCAGGATCCCGGCAGGCTGATCCCGTCCTCCACTGAGACCTG 539

Db     61 aaaaatggcttcggggcaaggcccaggtcctcccaggcaggagtgcggagagcctgccct 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    540 AAAAATGGCTTCGGGGCAAGGCCCAGGTCCTCCCAGGCAGGAGTGCGGAGAGCCTGCCCT 599

Db    121 gccctctgcttctgaggagcaggtagcccaggacacagaggaggttttccgcagctacgt 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    600 GCCCTCTGCTTCTGAGGAGCAGGTAGCCCAGGACACAGAGGAGGTTTTCCGCAGCTACGT 659

Db    181 tttttaccgccatcagcaggaacaggaggctgaaggggtggctgcccctgccgacccaga 240
          |||||||| |||||||||||||||||||||||||| ||||||||||||||||||||||||
Qy    660 TTTTTACCACCATCAGCAGGAACAGGAGGCTGAAGGGGCGGCTGCCCCTGCCGACCCAGA 719

Db    241 gatggtcaccttacctctgcaacctagcagcaccatggggcaggtgggacggcagctcgc 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    720 GATGGTCACCTTACCTCTGCAACCTAGCAGCACCATGGGGCAGGTGGGACGGCAGCTCGC 779

Db    301 catcatcggggacgacatcaaccgacgctatgactcagagttccagaccatgttgcagca 360
          |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    780 CATCATTGGGGACGACATCAACCGACGCTATGACTCAGAGTTCCAGACCATGTTGCAGCA 839

Db    361 cctgcagcccacggcagagaatgcctatgagtacttcaccaagattgccaccagcctgtt 420
          |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Qy    840 CCTGCAGCCCACGGCAGAGAATGCCTATGAGTACTTCACCAAGATTGCCTCCAGCCTGTT 899

Db    421 tgagagtggcatcaattggggccgtgtggtggctcttctgggcttcggctaccgtctggc 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    900 TGAGAGTGGCATCAATTGGGGCCGTGTGGTGGCTCTTCTGGGCTTCAGCTACCGTCTGGC 959

Db    481 cctacacgtctaccagcatggcctgactggcttcctaggccaggtgacccgcttcgtggt 540
          ||||||| |||||||| |||||||||||||||| |||||||||||||||| |||||
Qy    960 CCTACACATCTACCAGCGTGGCCTGACTGGCTTCCTGGGCCAGGTGACCCGCTTTGTGGT 1019

Db    541 cgacttcatgctgcatcactgcattgcccggtggattgcacagaggggtggctgggtggc 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1020 GGACTTCATGCTGCATCACTGCATTGCCCGGTGGATTGCACAGAGGGGTGGCTGGGTGGC 1079

Db    601 agccctgaacttgggcaatggtcccatcctgaacgtgctggtggttctgggtgtggttct 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1080 AGCCCTGAACTTGGGCAATGGTCCCATCCTGAACGTGCTGGTGGTTCTGGGTGTGGTTCT 1139

Db    661 gttgggccagtttgtggtacgaagattcttcaaatcatgactcccaagggtgccctttgg 720
          |||||||||||||||||||||||||||||||||||||||||||||||||||| || ||
Qy   1140 GTTGGGCCAGTTTGTGGTACGAAGATTCTTCAAATCATGACTCCCAAGGGTGCCTTTGGG 1199

Db    721 gtcccggttcagacccctgcctggacttaagcgaagtctttgccttctctgttcccttgc 780
          ||||| ||||||||||||||||||||||||||||||||||||||||||||| ||| ||||
Qy   1200 GTCCCAGTTCAGACCCCTGCCTGGACTTAAGCGAAGTCTTTGCCTTCTCTGCTCC-TTGC 1258

Db    781 aggggtccccccctcaagagtacagaagctt 810
          |||| ||||||||||||||||||||||||||
Qy   1259 AGGG-TCCCCCCTCAAGAGTACAGAAGCTT 1287


RESULT    6
ID   Q95494 standard; DNA; 5408 BP.
AC   Q95494;
DT   21-NOV-1995  (first entry)
DE   Human Cdn-3 DNA.
KW   Cdn-3; apoptosis modulator; adoptive immunotherapy; therapy; HIV;
KW   autoimmune disease; reperfusion injury; hepatitis, osteoporosis;
KW   shock; lymphoma; eczema; ss.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   CDS             1665..1931
FT   /*tag=  a
PN   WO9515084-A.
PD   08-JUN-1995.
PF   30-NOV-1994; U13930.
PR   30-NOV-1993; US-160067.
PR   07-OCT-1994; US-320157.
PA   (LXRB-) LXR BIOTECHNOLOGY INC.
PI   Barr PJ,  Kiefer MC;
DR   WPI; 95-215106/28.
DR   P-PSDB; R77878.
PT   New nucleic acid sequences encoding Cdn apoptosis modulators  - and
PT   related vectors, transformed cells, proteins and antibodies, useful
PT   or diagnosis and treatment e.g. of HIV infection, reperfusion injury
PT   etc.
PS   Disclosure; Fig.7A-G; 66pp; English.
CC   Southern blot analysis of human genomic DNA and a panel of
CC   human/rodent somatic cell DNAs revealed at least 3 Cdn-related
CC   genes residing on chromosomes 6, 11 and 20.  Cdn-3 (R77878)
CC   did not contain the structural features of Cdn-1 (R77876), Cdn-2
CC   (R77877) or other Bcl-2 family members.
SQ   Sequence   5408 BP;    1369 A;    1384 C;    1314 G;    1341 T;

  Query Match           58.5%;  Score 753;  DB 14;  Length 5408;
  Best Local Similarity  91.0%;  Pred. No. 0.00e+00;
  Matches   886;  Conservative    0;  Mismatches  79;  Indels   9;  Gaps   7;

Db   1438 atgatcactggagtctcacgggtccctcgggccacacagggacaagcaaaggctacatcc 1497
          |||||||||||||||| |||||||||| |||  |||||||||||| ||||||||||||
Qy    317 ATGATCACTGGAGTCTCGCGGGTCCCTCAGGCTGCACAGGGACAAGTAAAGGCTACATCC 376

Db   1498 agataccagaaatgcactgacgcccgttcctggaagctgggctcccactcagcccctggg 1557
          |||| |  | |||||||||||||| ||||||||| ||||||||||||||||||||||||
Qy    377 AGATGCTGGGAATGCACTGACGCCCATTCCTGGAAACTGGGCTCCCACTCAGCCCCTGGG 436

Db   1558 agcagcagcctccagccccttgggaccttcaactccaccctgctgacccacgcgggttga 1617
          |||||||||| ||||||||| ||||||| || |||||||||||||| |||| |||||||
Qy    437 AGCAGCAGCCGCCAGCCCCTCGGGACCTCCATCTCCACCCTGCTGAGCCACCCGGGTTGG 496

Db   1618 gccagcatccctggaggctgacactgtcctccactgagacctgaaaaatggcatcggggc 1677
          ||||| ||||| | |||||||  | ||||||||||||||||||||||||||| |||||||
```

```
Qy    497 GCCAGGATCCCGGCAGGCTGATCCCGTCCTCCACTGAGACCTGAAAAATGGCTTCGGGGC 556

Db   1678 aaggcccagggcctcccaggcaggagtgcggaaagcctgccctgccctctgcttctgagg 1737
          ||||||||| |||||||||||||||||||| |||||||||||||||||||||||||||||
Qy    557 AAGGCCCAGGTCCTCCCAGGCAGGAGTGCGGAGAGCCTGCCCTGCCCTCTGCTTCTGAGG 616

Db   1738 agcaggtagcccaggacatggagg-ggttttccgcagctacgtttttttaccaccatcagc 1796
          |||||||||||||||||  |||| ||||||||||||||||||||||||||||||||||||
Qy    617 AGCAGGTAGCCCAGGACACAGAGGAGGTTTTCCGCAGCTACGTTTTTTACCACCATCAGC 676

Db   1797 aggaacaggaggctgaaggggcggccgcccctgccgacccagagatggtcaccttgcccc 1856
          |||||||||||||||||||||||| |||||||||||||||||||||||||||||| || |
Qy    677 AGGAACAGGAGGCTGAAGGGGCGGCTGCCCCTGCCGACCCAGAGATGGTCACCTTACCTC 736

Db   1857 tccaacctagcagcaccatggggcaggtgggacggcagctcgccatcaccagg-acgaca 1915
          | ||||||||||||||||||||||||||||||||||||||||||||||||   || ||||||
Qy    737 TGCAACCTAGCAGCACCATGGGGCAGGTGGGACGGCAGCTCGCCATCATTGGGGACGACA 796

Db   1916 tcaaccggcactatgacttcggagttccagaccatgctgcagcacctgcagcccacggca 1975
          ||||||| | |||||||| | |||||||||||||| ||||||||||||||||||||||||
Qy    797 TCAACCGACGCTATGACT-CAGAGTTCCAGACCATGTTGCAGCACCTGCAGCCCACGGCA 855

Db   1976 gagaacgcctacgagtacttcaccaagatcgcctccagcctgtttgagagtggcatcaac 2035
          ||||| ||||| |||||||||||||||| |||||||||||||||||||||||||||||||
Qy    856 GAGAATGCCTATGAGTACTTCACCAAGATTGCCTCCAGCCTGTTTGAGAGTGGCATCAAT 915

Db   2036 cggggccgtgtggtggctctcctgggcttcggctaccgtctggtcctacatgtctaccag 2095
           |||||||||||||||||| |||||||| |||||||||||| ||||||  ||||||||
Qy    916 TGGGGCCGTGTGGTGGCTCTTCTGGGCTTCAGCTACCGTCTGGCCCTACACATCTACCAG 975

Db   2096 cacggcttgactggcttcctgggcctggtgacccgcttcgtggt---cttcatgctgcaa 2152
          |  ||| |||||||||||||||||| |||||||||||| |||||   ||||||||||||
Qy    976 CGTGGCCTGACTGGCTTCCTGGGCCAGGTGACCCGCTTTGTGGTGGACTTCATGCTGCAT 1035

Db   2153 caaggcatcgcccggtggatctcgcagaggggcggctgggtggcagccctggacttgggc 2212
          ||  |||| ||||||||||  | |||||||| |||||||||||||||||| ||||||||
Qy   1036 CACTGCATTGCCCGGTGGATTGCACAGAGGGGTGGCTGGGTGGCAGCCCTGAACTTGGGC 1095

Db   2213 aatagtcccatcctgaacgtgctggtggttgtgggtgtggttctgctgggccagtttgtg 2272
          ||| |||||||||||||||||||||||||| |||||||||||||| ||||||||||||||
Qy   1096 AATGGTCCCATCCTGAACGTGCTGGTGGTTCTGGGTGTGGTTCTGTTGGGCCAGTTTGTG 1155

Db   2273 gtaagaagattcttcaaatcatgactcccaggggtgtccttttggggtcccagctgtgacc 2332
          ||| |||||||||||||||||||||||||| ||||| ||||||||||||||| |  ||||
Qy   1156 GTACGAAGATTCTTCAAATCATGACTCCCAAGGGTG-CCTTTGGGGTCCCAGTTCAGACC 1214

Db   2333 cctgcctggacttaagccaagtctttgccttccccactcccttgcaggggtcaccccttca 2392
          |||||||||||||||| |||||||||||||| |  |||| |||||||| || ||| |||
Qy   1215 CCTGCCTGGACTTAAGCGAAGTCTTTGCCTTCTCTGCTCC-TTGCAGGG-TCCCCCCTCA 1272

Db   2393 aaagtacagaagct 2406
          | ||||||||||||
Qy   1273 AGAGTACAGAAGCT 1286


RESULT    7
ID   Q10572 standard; DNA; 1047 BP.
AC   Q10572;
DT   09-APR-1991  (first entry)
DE   Human Natriuretic Peptide Receptor B.
KW   NPRB; ANP; BNP; CNP; kidney failure; heart failure; protein kinase;
KW   hyperaldosteronism; glaucoma; guanyl cyclase.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   Peptide         1..22
FT   /label= signal sequence
FT   Protein         12
FT   /label= mature NPBR
FT   Domain          23..455
FT   /label= extracellular domain
FT   /note= "binds natriuretic peptides A,B and C]"
FT   Domain          456..456
FT   /label= transmembrane domain
FT   Domain          479..1047
FT   /label= cytoplasmic domain
FT   /note= "GC and protien kinase activity"
FT   Modified -site  24..26
FT   /label= N-glycos_site
FT   Modified -site  35..37
FT   /label= N-glycos_site
FT   Modified -site  161..163
FT   /label= N-glycos_site
FT   Modified -site  195..197
FT   /label= N-glycos_site
FT   Modified -site  244..246
FT   /label= N-glycos_site
FT   Modified -site  277..279
FT   /label= N-glycos_site
FT   Modified -site  349..351
FT   /label= N-glycos_site
FT   Modified -site  600..602
FT   /label= N-glycos_site
PN   WO9100292-A.
PD   10-JAN-1991.
PF   22-JUN-1990; U03586.
PR   23-JUN-1989; US-370673.
PA   (GETH ) GENENTECH INC.
PI   Chang M, Goeddel D, Lowe D;
DR   WPI; 91-036711/05.
DR   N-PSDB; Q10324.
PT   Natriuretic protein receptor B - for diagnosis and treatment of
PT   kidney failure, heart failure, hyperaldosteronism, glaucoma etc.
PS   Claim 3; Fig 1; 49pp; English.
CC   The sequence was derived from the DNA encoding natriuretic peptide
CC   receptor B, NPRB, having guanyl cyclase (GC) activity and protein
CC   kinase activity.  The DNA can be inserted into expression vectors
CC   for the prodn. of the protein, opt. after being mutated to produce
CC   NPRB analogues. The protein has a mol wt. of 115 kD (calculated Mr=
CC   114,952).  The protein (or variants) can be used in treatment of
CC   natriuretic peptide disorders, and also to isolate peptides using
CC   affinity chromatography.  Antibodies with affinity for NPRB can
CC   also be prepd.
SQ   Sequence   1047 BP;    87 A;    15 C;    83 G;    51 T;

  Query Match             6.8%;  Score 88;  DB 2;  Length 1047;
  Best Local Similarity   8.3%;  Pred. No. 5.85e-38;
  Matches    61;  Conservative  215;  Mismatches 450;  Indels   6;  Gaps   6;

Db    299 rnnnnnnynnnnnrnnnrarndngvnngnsnmnnnagcnydgnnnyanvnnntnnnggtr 358
          :          : : : |:    : :   |||  :|      :       | |
Cp   1287 AAGCTTCTGTACTCTTGAGGGGGGACCCTGCAAGGAGCAGAGAAGGCAAAGACTTCGCTT 1228

Db    359 ndgnrnvnkmngrryhgvtgnvvmdknndrntdnvnwamgdndsgdnnnaahysganknn 418
           :|   : :: |:: : :|| :: ::      : :   :   : |:     ::: | :
Cp   1227 AAGTCCAGGCAGGGGT-CTGAACTGGGACCCCAAAGGCACCCTTGGGAGTCATG-ATTTG 1170

Db    419 wwtgrnnnwvkgannsdnnncandnddnscdktnnstnanvangtgntnnmngvssnnnn 478
          ::  :   ::: |       |  :     |::   :  | :  |       : |: :
Cp   1169 AAGAATCTTCGTACCACAAACTGGCCCAACAGAACCACACCCAGAACCACCAGCACGTTC 1110

Db    479 rknmnnknnasmnwrnrwnnnnngnsnryhkgagsrntnsnrgssygsnmtahgkynnna 538
          :: :  :   :: ::          : ::::  :    : : ::: :  | :  :
Cp   1109 AGGATGGGACCATTGCCCAAGTTCAGGGCTGCCACCCAGCCACCCCTCTGTGCAATCCAC 1050

Db    539 ntghnkgnvvankhvnkkrnnntrnvnnnnkhmrdvnnnhntrnngacndnnnncnvtny 598
          |     ::| : : : :   |: :     ::: :   :  :    | :      :
Cp   1049 CGGGCAATGCAGTGATGCAGCATGAAGTCCACCACAAAGCGGGTCACCTGGCCCAGGAAG 990

Db    599 cnrgsnndnnnndsnnndwmnrysnnndnvkgmannhnsnnsshgsnkssncvvdsrnvn 658
          | :|   :      :   ::: :::   : :: :|  : :  :::|  :::  ::::: :
Cp    989 CCAGTCAGGCCACGCTGGTAGATGTGTAGGGCCAGACGGTAGCTGAAGCCCAGAAGAGCC 930

Db    659 kntdygnasnrstannddnnanyakknntannnnsgnnnnttgmnaadvysngnnnnnna 718
             ::   :  : |  ::    :| :   |    :|      |: ||::  |      |
Cp    929 ACCACACGGCCCCAATTGATGCCACTCTCAAACAGGCTGGAGGC-AATCTTGGTGAAGTA 871
```

```
Db    719 nrsgnnynngndnsnknnvnkvrngnrnynrnsndrtnnnnnnvnnmnrcwandnanrnd 778
Cp    870 CTCATAGGCATTCTCTGCCGTGGGCTGCAGGTGCTG-CAACATGGTCTGGAACTCTGAGT 812

Db    779 ngnnkgnnrrnnknggtsnndnnnnrmnnyannnnknvnnrtnaynnnkrkanannynnn 838
Cp    811 CATAGCGTCGGTT-GATGTCGTCCCCAATGATGGCGAGCTGCCGTCCCACCTGCCCCATG 753

Db    839 nhsvannnkrgntvnanandsvtnynsdnvgntansanstnmnvvtnnndnytcndannd 898
Cp    752 GTGCTGCTAGGTTGCAGAGGTAAGGTGACCATCTCTGGGTCGGCAGGGGCAGCCGCCCCT 693

Db    899 nndvykvntngdaymvvsgnngrngnrhannnarmananndavssnrnrhrnh-dnnrnr 957
Cp    692 TCAGCCTCCTGTTCCTGCTGATGGTGGTAAAAAACGTAGCTGCGGAAAACCTCCTCTGTG 633

Db    958 ngvhtgnvcagvvgnkmnrycnngdtvntasrmnsngnanknhvssttkdandnngcnnn 1017
Cp    632 TCCTGGGCTACCTGCTCCTCAGAAGCAGAGGGCAGGGCAGGCTCTCCGCACTCCTGCCTG 573

Db   1018 nnrgdvnmkgkg 1029
Cp    572 GGAGGACCTGGG 561
```

```
RESULT    8
ID   Q10572 standard; DNA; 1047 BP.
AC   Q10572;
DT   09-APR-1991  (first entry)
DE   Human Natriuretic Peptide Receptor B.
KW   NPRB; ANP; BNP; CNP; kidney failure; heart failure; protein kinase;
KW   hyperaldosteronism; glaucoma; guanyl cyclase.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   Peptide         1..22
FT   /label= signal sequence
FT   Protein         12
FT   /label= mature NPBR
FT   Domain          23..455
FT   /label= extracellular domain
FT   /note= "binds natriuretic peptides A,B and C]"
FT   Domain          456..456
FT   /label= transmembrane domain
FT   Domain          479..1047
FT   /label= cytoplasmic domain
FT   /note= "GC and protien kinase activity"
FT   Modified -site  24..26
FT   /label= N-glycos_site
FT   Modified -site  35..37
FT   /label= N-glycos_site
FT   Modified -site  161..163
FT   /label= N-glycos_site
FT   Modified -site  195..197
FT   /label= N-glycos_site
FT   Modified -site  244..246
FT   /label= N-glycos_site
FT   Modified -site  277..279
FT   /label= N-glycos_site
FT   Modified -site  349..351
FT   /label= N-glycos_site
FT   Modified -site  600..602
FT   /label= N-glycos_site
PN   WO9100292-A.
PD   10-JAN-1991.
PF   22-JUN-1990; U03586.
PR   23-JUN-1989; US-370673.
PA   (GETH ) GENENTECH INC.
PI   Chang M, Goeddel D, Lowe D;
DR   WPI; 91-036711/05.
DR   N-PSDB; Q10324.
PT   Natriuretic protein receptor B - for diagnosis and treatment of
PT   kidney failure, heart failure, hyperaldosteronism, glaucoma etc.
PS   Claim 3; Fig 1; 49pp; English.
```

```
CC   The sequence was derived from the DNA encoding natriuretic peptide
CC   receptor B, NPRB, having guanyl cyclase (GC) activity and protein
CC   kinase activity.  The DNA can be inserted into expression vectors
CC   for the prodn. of the protein, opt. after being mutated to produce
CC   NPRB analogues. The protein has a mol wt. of 115 kD (calculated Mr=
CC   114,952).  The protein (or variants) can be used in treatment of
CC   natriuretic peptide disorders, and also to isolate peptides using
CC   affinity chromatography.  Antibodies with affinity for NPRB can
CC   also be prepd.
SQ   Sequence   1047 BP;    87 A;    15 C;    83 G;    51 T;

  Query Match           6.0%;  Score 77;  DB 2;  Length 1047;
  Best Local Similarity   9.3%;  Pred. No. 5.42e-31;
  Matches    63;  Conservative  199;  Mismatches 412;  Indels   7;  Gaps   6;

Db    351 tnnnggtrndgnrnvnkmngrryhgvtgnvvmdknndrntdnvnwamgdndsgdnnnaah 410
Qy    185 TTTTGGTGTATGAACTGTAGTCCTAGAGGATTTTATTAGTTATGAGTTCTATAACTAAGA 244

Db    411 ysganknnwwtgrnnn-wvkgannsdnnncandnddnscdktnnstnanvangtgntnnm 469
Qy    245 TCCATCATCTTAGTTGCTAAGAACGTAGATACTGAGAACATCATTTAAAAAAACATTTTT 304

Db    470 ngvssnnnnrknmnnknnasmnwrnrwnnnnngnsnryhkgagsrntnsnrgssygsnmt 529
Qy    305 GGCTGGCACCTCATGATCACTG-GAGTCTCGCGGGTCCCTCAGGCTGCACAGGGACAAGT 363

Db    530 ahgkynnnantghnkgnvvankhvnkkrnnntrnvnnnnkhmrdvnnnhntrnngacndn 589
Qy    364 AAAGGCTACATCCAGATGCTGGGAATGCACTGACGCCCATTCCTGGAAACTGGGCTCCCA 423

Db    590 nnncnvtnycnrg-snndnnnndsnnndwmnrysnnndnvkgmannhnsnnsshgsnkss 648
Qy    424 CTCAGCCCCTGGGAGCAGCAGCCGCCAGCCCCTCGGGACCTCCATCTCCACCCTGCTGAG 483

Db    649 ncvvdsrnvnkntdygnasnrstannddnnanyakknntannnnsgnnnnttgmn--aad 706
Qy    484 CCACCCGGGTTGGGCCAGGATCCCGGCAGGCTGATCCCGTCCTCCACTGAGACCTGAAAA 543

Db    707 vysngnnnnnnanrsgnnynngndnsnknnvnkvrngnrnynrnsndrtnnnnnnvnnmn 766
Qy    544 ATGGCTTCGGGGCAAGGCCCAGGTCCTCCCAGGCAGGAGTGCGGAGAGCCTGCCCTGCCC 603

Db    767 rcwandnanrndngnnkgnnrrnnknggtsnndnnnnrmnnyannnnknvnnrtnaynnn 826
Qy    604 TCTGCTTCTGAGGAGCAGGTAGCCCAGGACACAGAGGAGGTTTTCCGCAGCTACGTTTTT 663

Db    827 krkanannynnnnhsvannnkrgntvnanandsvtnynsdnvgntansanstnmnvvtnn 886
Qy    664 TACCACCATCAGCAGGAACAGGAGGCTGAAGGGGCGGCTGCCCCTGCCGACCCAGAGATG 723

Db    887 ndnytcndanndnndvykvntngdaymvvsgnngrngnrhannnarmananndavssnrn 946
Qy    724 GTCACCTTACCTCTGCAACCTAGCAGCACCATGGGGCAGGTGGGACGGCAGCT-CGCCAT 782

Db    947 rhrnhdnnrnrngvhtgnvcagvvgnkmnrycnngdtvntasrmnsngnanknhvssttk 1006
Qy    783 -CATTGGGGACGACATCAACCGACGCTATGACTCAGAGTTCCAGACCATGTTGCAGCACC 841

Db   1007 dandnngcnnnnnrgdvnmkg 1027
Qy    842 TGCAGCCCACGGCAGAGAATG 862
```

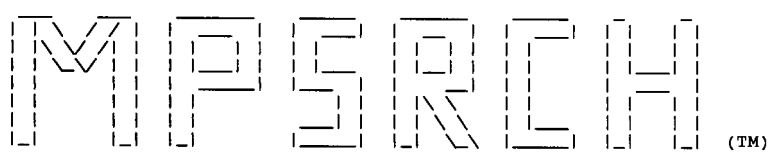
```
RESULT    9
ID   N81164 standard; DNA; 204 BP.
AC   N81164;
DT   08-NOV-1990  (first entry)
DE   Base substituted E.coli beta-galactosidase alpha-fragment.
KW   E.coli beta galactosidase alpha-fragment; base substitutions; ss.
OS   Escherichia coli.
FH   Key             Location/Qualifiers
FT   misc_feature    19..69
```

```
FT   /*tag= a
FT   /function=multiple cloning site
FT   primer_bind     187..204
FT   /*tag= b
PN   EP-285123-A.
PD   05-MAY-1988.
PF   30-MAR-1988; 105163.
PR   03-APR-1987; US-034819.
PA   (SUSO) SUOMEN SOKERI OY.
PI   Lehtovaara P, Knowles J, Koivula A, Bamford J, Reinikainen T;
DR   WPI; 88-279927/40.
PT   Introducing random point mutations into nucleic acods -
PT   by prepn of single stranded template, annealing a primer, elongation,
PT   misincorporation, completion of molecules and screening.
PS   Disclosure; p; English.
CC   Random point mutations were introduced into the alpha fragment of
CC   E.coli beta-galactosidase. The wild type sequence was obtained as a
CC   single stranded template and an oligonucleotide was hybridised to
CC   it to generate a popn of DNA molecules which terminate at all
CC   possible nucleotide positions within a specified region. The
CC   variable 3' ends generated in this way are used as primers for
CC   reverse transcriptase. Nucleotides are misincorporated by the
CC   transcriptase and the molecules are completed to forms that can be
CC   amplified and then expressed in a suitable host-vector system.
CC   The sequence covers all 176 difft base substitutions, most of which
CC   occurred singularly in any given mutant.
CC   See also P80575.
SQ   Sequence   204 BP;   21 A;   47 C;   17 G;   11 T;   108 Others;

  Query Match            3.5%;  Score 45;  DB 1;  Length 204;
  Best Local Similarity  11.2%;  Pred. No. 1.24e-11;
  Matches    12;  Conservative   57;  Mismatches  38;  Indels   0;  Gaps   0;

Db     86 ymrtttthhyrrmrbnvyrdynrsdaaawyccyrrsvkydccynachhddhyvybbbvynv 145
          ::: | :: : ::: :: :  :::   ::||::: ::::||:     ::::::: : : :
Cp   1228 TAAGTCCAGGCAGGGGTCTGAACTGGGACCCCAAAGGCACCCTTGGGAGTCATGATTTGA 1169

Db    146 hnhnncncccbnnhvchnvhbnnhrnwayvrhdarrddvhccvchcc 192
          : :  |  | :  ::|: :::  :  :|:::::  : :::| : :||
Cp   1168 AGAATCTTCGTACCACAAACTGGCCCAACAGAACCACACCCAGAACC 1122


RESULT   10
ID   Q51746 standard; cDNA; 91 BP.
AC   Q51746;
DT   31-MAY-1994  (first entry)
DE   Oligonucleotide probe MK14-A
KW   Oligonucleotide; DNA probe; mycobacteria; disease diagnosis;
KW   ss.
OS   Synthetic.
PN   EP-571911-A.
PD   01-DEC-1993.
PF   24-MAY-1993; 108325.
PR   26-MAY-1992; US-889651.
PA   (BECT ) BECTON DICKINSON CO.
PI   Shank DD,  Spears PA;
DR   WPI; 93-378844/48.
PT   New oligo:nucleotide probes specific for Mycobacteria - used for
PT   detection and amplification of Mycobacteria nucleic acid in
PT   samples
PS   Claim 3; Page 14; 23pp; English.
CC   Oligonucleotide probe MK14-A consists of nucleotides 5-95 of MK14
CC   (Q51735).  It hybridized to all spp. of mycobacteria tested, but
CC   cross reacted to a few non-mycobacterial spp.  The probe may
CC   be useful as an initial screen for mycobacterial infection.
CC   See also Q51735-45 and Q51747-59.
SQ   Sequence   91 BP;    5 A;    17 C;    15 G;    4 T;

  Query Match            3.3%;  Score 43;  DB 9;  Length 91;
  Best Local Similarity  10.2%;  Pred. No. 1.70e-10;
  Matches     6;  Conservative   45;  Mismatches   8;  Indels   0;  Gaps   0;

Db      3 ctccggcgcssvhsyyvvhvvshhhsvhhvvhhvhvsvvvvhhvvhvvhhvhyhvyvsvc 61
          |||| | :::::::::::::::::::: ::::: ::::::::::::::::: :::  ::|
Qy    463 CTCCATCTCCACCCTGCTGAGCCACCCGGGTTGGGCCAGGATCCCGGCAGGCTGATCCC 521


RESULT   11
ID   Q51746 standard; cDNA; 91 BP.
AC   Q51746;
DT   31-MAY-1994  (first entry)
DE   Oligonucleotide probe MK14-A
KW   Oligonucleotide; DNA probe; mycobacteria; disease diagnosis;
KW   ss.
OS   Synthetic.
PN   EP-571911-A.
PD   01-DEC-1993.
PF   24-MAY-1993; 108325.
PR   26-MAY-1992; US-889651.
PA   (BECT ) BECTON DICKINSON CO.
PI   Shank DD,  Spears PA;
DR   WPI; 93-378844/48.
PT   New oligo:nucleotide probes specific for Mycobacteria - used for
PT   detection and amplification of Mycobacteria nucleic acid in
PT   samples
PS   Claim 3; Page 14; 23pp; English.
CC   Oligonucleotide probe MK14-A consists of nucleotides 5-95 of MK14
CC   (Q51735).  It hybridized to all spp. of mycobacteria tested, but
CC   cross reacted to a few non-mycobacterial spp.  The probe may
CC   be useful as an initial screen for mycobacterial infection.
CC   See also Q51735-45 and Q51747-59.
SQ   Sequence   91 BP;    5 A;    17 C;    15 G;    4 T;

  Query Match            3.3%;  Score 42;  DB 9;  Length 91;
  Best Local Similarity  0.0%;  Pred. No. 6.22e-10;
  Matches     0;  Conservative   46;  Mismatches   4;  Indels   0;  Gaps   0;

Db     11 ssvhsyyvvhvvshhhsvhhvvhhvhvsvvvvhhvvhvvhhvhyhvyvsv 60
          ::::::::::::::::::::::: ::::  :: ::::::::::::::::::
Cp   1204 GGGACCCCAAAGGCACCCTTGGGAGTCATGATTTGAAGAATCTTCGTACC 1155


RESULT   12
ID   N81164 standard; DNA; 204 BP.
AC   N81164;
DT   08-NOV-1990  (first entry)
DE   Base substituted E.coli beta-galactosidase alpha-fragment.
KW   E.coli beta galactosidase alpha-fragment; base substitutions; ss.
OS   Escherichia coli.
FH   Key             Location/Qualifiers
FT   misc_feature    19..69
FT   /*tag= a
FT   /function=multiple cloning site
FT   primer_bind     187..204
FT   /*tag= b
PN   EP-285123-A.
PD   05-MAY-1988.
PF   30-MAR-1988; 105163.
PR   03-APR-1987; US-034819.
PA   (SUSO) SUOMEN SOKERI OY.
PI   Lehtovaara P, Knowles J, Koivula A, Bamford J, Reinikainen T;
DR   WPI; 88-279927/40.
PT   Introducing random point mutations into nucleic acods -
PT   by prepn of single stranded template, annealing a primer, elongation,
PT   misincorporation, completion of molecules and screening.
PS   Disclosure; p; English.
CC   Random point mutations were introduced into the alpha fragment of
CC   E.coli beta-galactosidase. The wild type sequence was obtained as a
CC   single stranded template and an oligonucleotide was hybridised to
CC   it to generate a popn of DNA molecules which terminate at all
CC   possible nucleotide positions within a specified region. The
CC   variable 3' ends generated in this way are used as primers for
CC   reverse transcriptase. Nucleotides are misincorporated by the
CC   transcriptase and the molecules are completed to forms that can be
CC   amplified and then expressed in a suitable host-vector system.
CC   The sequence covers all 176 difft base substitutions, most of which
```

```
CC   occurred singularly in any given mutant.
CC   See also P80575.
SQ   Sequence   204 BP;    21 A;    47 C;    17 G;    11 T;    108 Others;

  Query Match              3.1%;  Score 40;  DB 1;  Length 204;
  Best Local Similarity   12.9%;  Pred. No. 8.16e-09;
  Matches    13;  Conservative    53;  Mismatches  34;  Indels    1;  Gaps    1;

Db       95 rrmrbnvyrdynrsdaaawyccyrrsvkydccynachhddhyvybbbvynvhnhnncncc 154
            ::::: :: :   : :|| ::   : :::: :| :   |:::::::::::   ::        |
Qy      281 AACATCATTTAAAAAAACATTTTGGCTGGCACCTCATGATCACTGGAGTCTCGCGGGTC 340

Db      155 cbnnhvchnvhbnnhrnwayvrhdarrddvhccvchccgat 195
            |:  :: : ::    ::  |::: :|:::::: |: :| |||
Qy      341 CCTCAGGCTGC-ACAGGGACAAGTAAAGGCTACATCCAGAT 380


RESULT   13
ID   Q70468 standard; DNA; 114 BP.
AC   Q70468;
DT   05-APR-1995  (first entry)
DE   Generic DNA sequence to generate a random TSAR petide library.
KW   TSAR; totally synthetic affinity reagent; synthetic; binding domain;
KW   effector domain; concateneated heterofunctional protein; linker;
KW   direct; rapid; detection; screening; treatment; generic; ss.
OS   Synthetic.
FH   Key             Location/Qualifiers
FT   misc_feature    55..60
FT   /*tag=  a
FT   /note= "this sequence represents 'Z'; Z can be a
FT   sequence of 6, 9 or 12 nucleotides (see
FT   comments)"
PN   WO9418318-A.
PD   18-AUG-1994.
PF   01-FEB-1994; U00977.
PR   01-FEB-1993; US-013416.
PR   30-DEC-1993; US-176500.
PR   31-JAN-1994; US-189331.
PA   (UYNC-) UNIV NORTH CAROLINA.
PI   Fowlkes DM,  Kay BK;
DR   WPI; 94-279739/34.
DR   P-PSDB; R65154.
PT   Identifying proteins or peptide(s) which bind a ligand - by
PT   screening a recombinant vector library expressing fusion proteins
PT   comprising a binding domain and an effector domain
PS   Disclosure; Page 35; 255pp; English.
CC   Q70468 is a generic DNA sequence used to generate random TSAR (Totally
CC   Synthetic Affinity Reagents) peptides.This generic formula can also be
CC   represented  as follows: X(NNB)11(TGC)(NNB)6Z(NNB)7(TGC)(NNB)10Y. X
CC   and Y are flanking  restriction sites (X is not the same as Y) that are
CC   not specified further. Other generic sequences are shown in Q70466-68.
CC   Other specific peptides generated by these generic sequences are shown in
CC   R65151-54. TSARs are concatenated heterofunctional proteins or peptides,
CC   comprising at least two functional regions - a binding domain with
CC   affinity for a ligand and a second effector peptide portion that is
CC   chemically or biologically active.They may further comprise a linker
CC   peptide between the 2 domains.The oligonucleotides are also designed so
CC   that the expressed peptide contains 2 or 4 cysteine residues positioned
CC   in, or flanking, the unpredicted or variant residues. These residues
CC   confer some degree of conformational rigidity to the peptides. The TSARs
CC   or compsns. comprising a TSAR binding domain can be used in vivo to
CC   deliver a chemically or biologically active moiety, eg. metal ion,
CC   radioisotope, peptide, toxin or enzyme, to the specific target or on the
CC   cell. They can also replace the function of macromolecules, eg.
CC   monoclonal or polyclonal antibodies and therefore circumvent the need
CC   for complex methods of hybridoma formation or in vivo antibody
CC   production. The TSARs are easily characterised and have designed activity
CC   allowing direct and rapid detection in a screening process.
SQ   Sequence   114 BP;    0 A;    2 C;    2 G;    2 T;

  Query Match              2.8%;  Score 36;  DB 12;  Length 114;
  Best Local Similarity    3.6%;  Pred. No. 1.26e-06;
  Matches     4;  Conservative    34;  Mismatches  74;  Indels    0;  Gaps    0;

Db        3 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbtgcnnbnnbnnbnnbnnbnnbnnnnnnnn 62
              :  :  :  :  :  :  :  :  :  :  : ||  :  :  :  :  :  :
Qy      909 CATCAATTGGGGCCGTGTGGTGGCTCTTCTGGGCTTCAGCTACCGTCTGGCCCTACACAT 968

Db       63 bnnbnnbnnbnnbnnbnnbtgcnnbnnbnnbnnbnnbnnbnnbnnbnnb 114
              :  :  :  :  :  :  ||  :  :  :  :  :  :  :  :  :
Qy      969 CTACCAGCGTGGCCTGACTGGCTTCCTGGGCCAGGTGACCCGCTTTGTGGTG 1020


RESULT   14
ID   Q70466 standard; DNA; 114 BP.
AC   Q70466;
DT   05-APR-1995  (first entry)
DE   Generic DNA sequence to generate a random TSAR-9 petide library.
KW   TSAR; totally synthetic affinity reagent; synthetic; binding domain;
KW   effector domain; concateneated heterofunctional protein; linker;
KW   direct; rapid; detection; screening; treatment; generic; ss.
OS   Synthetic.
FH   Key             Location/Qualifiers
FT   misc_feature    55..60
FT   /*tag=  a
FT   /note= "this sequence represents 'Z'; Z can be a
FT   sequence of 6, 9 or 12 nucleotides (see
FT   comments)"
PN   WO9418318-A.
PD   18-AUG-1994.
PF   01-FEB-1994; U00977.
PR   01-FEB-1993; US-013416.
PR   30-DEC-1993; US-176500.
PR   31-JAN-1994; US-189331.
PA   (UYNC-) UNIV NORTH CAROLINA.
PI   Fowlkes DM,  Kay BK;
DR   WPI; 94-279739/34.
DR   P-PSDB; R65152.
PT   Identifying proteins or peptide(s) which bind a ligand - by
PT   screening a recombinant vector library expressing fusion proteins
PT   comprising a binding domain and an effector domain
PS   Disclosure; Page 35; 255pp; English.
CC   Q70466 is a generic DNA sequence used to generate random TSAR (Totally
CC   Synthetic Affinity Reagents) peptides.This generic formula can also be
CC   represented as follows:X(NNB)1(TGC)(NNB)10(TGC)2(NNB)4Z(NNB)8(TGC)(NNB)
CC   -9Y. X and Y are flanking  restriction sites (X is not the same as Y)
CC   that are not specified further. Other generic sequences are shown in
CC   Q70466-68. Other specific peptides generated by these generic sequences
CC   are shown in R65151-54. TSARs are concatenated heterofunctional proteins
CC   or peptides, comprising at least two functional regions - a binding
CC   domain with affinity for a ligand and a second effector peptide portion
CC   that is chemically or biologically active.They may further comprise a
CC   linker peptide between the 2 domains.The oligonucleotides are also
CC   designed so that the expressed peptide contains 2 or 4 cysteine residues
CC   positioned in, or flanking, the unpredicted or variant residues. These
CC   residues confer some degree of conformational rigidity to the peptides.
CC   The TSARs or compsns. comprising a TSAR binding domain can be used in
CC   vivo to deliver a chemically or biologically active moiety, eg. metal
CC   ion, radioisotope, peptide, toxin or enzyme, to the specific target or
CC   on the cell. They can also replace the function of macromolecules, eg.
CC   monoclonal or polyclonal antibodies and therefore circumvent the need for
CC   complex methods of hybridoma formation or in vivo antibody production.
CC   The TSARs are easily characterised and have designed activity allowing
CC   direct and rapid detection in a screening process.
SQ   Sequence   114 BP;    0 A;    4 C;    4 G;    4 T;

  Query Match              2.8%;  Score 36;  DB 12;  Length 114;
  Best Local Similarity    7.3%;  Pred. No. 1.26e-06;
  Matches     8;  Conservative    31;  Mismatches  71;  Indels    0;  Gaps    0;

Db        5 gcnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbtgctgcnnbnnbnnbnnbnnnnnnnnbn 64
            ||  :  :  :  :  :  :  :  :  :  : ||| |  :  :  :  :          :
Qy      908 GCATCAATTGGGGCCGTGTGGTGGCTCTTCTGGGCTTCAGCTACCGTCTGGCCCTACACA 967

Db       65 nbnnbnnbnnbnnbnnbnnbtgcnnbnnbnnbnnbnnbnnbnnbnnbnnb 114
              :  :  :  :  :  :  ||  :  :  :  :  :  :  :  :  :
```

```
Qy    968 TCTACCAGCGTGGCCTGACTGGCTTCCTGGGCCAGGTGACCCGCTTTGTG 1017


RESULT   15
ID   Q70467 standard; DNA; 114 BP.
AC   Q70467;
DT   05-APR-1995  (first entry)
DE   Generic DNA sequence to generate a random TSAR petide library.
KW   TSAR; totally synthetic affinity reagent; synthetic; binding domain;
KW   effector domain; concatenated heterofunctional protein; linker;
KW   direct; rapid; detection; screening; treatment; generic; ss.
OS   Synthetic.
FH   Key             Location/Qualifiers
FT   misc_feature    55..60
FT   /*tag= a
FT   /note= "this sequence represents 'Z'; Z can be a
FT   sequence of 6, 9 or 12 nucleotides (see
FT   comments)"
PN   WO9418318-A.
PD   18-AUG-1994.
PF   01-FEB-1994; U00977.
PR   01-FEB-1993; US-013416.
PR   30-DEC-1993; US-176500.
PR   31-JAN-1994; US-189331.
PA   (UYNC-) UNIV NORTH CAROLINA.
PI   Fowlkes DM,  Kay BK;
DR   WPI; 94-279739/34.
DR   P-PSDB; R65153.
PT   Identifying proteins or peptide(s) which bind a ligand - by
PT   screening a recombinant vector library expressing fusion proteins
PT   comprising a binding domain and an effector domain
PS   Disclosure; Page 35; 255pp; English.
CC   Q70467 is a generic DNA sequence used to generate random TSAR (Totally
CC   Synthetic Affinity Reagents) peptides.This generic formula can also be
CC   represented  as follows: X(NNB)16(TGC)(NNB)1Z(NNB)16(TGC)(NNB)1Y. X
CC   and Y are flanking  restriction sites (X is not the same as Y) that are
CC   not specified further. Other generic sequences are shown in Q70466-68.
CC   Other specific peptides generated by these generic sequences are shown in
CC   R65151-54. TSARs are concatenated heterofunctional proteins or peptides,
CC   comprising at least two functional regions - a binding domain with
CC   affinity for a ligand and a second effector peptide portion that is
CC   chemically or biologically active.They may further comprise a linker
CC   peptide between the 2 domains.The oligonucleotides are also designed so
CC   that the expressed peptide contains 2 or 4 cysteine residues positioned
CC   in, or flanking, the unpredicted or variant residues. These residues
CC   confer some degree of conformational rigidity to the peptides. The TSARs
CC   or compsns. comprising a TSAR binding domain can be used in vivo to
CC   deliver a chemically or biologically active moiety, eg. metal ion,
CC   radioisotope, peptide, toxin or enzyme, to the specific target or on the
CC   cell. They can also replace the function of macromolecules, eg.
CC   monoclonal or polyclonal antibodies and therefore circumvent the need for
CC   complex methods of hybridoma formation or in vivo antibody production.
CC   The TSARs are easily characterised and have designed activity allowing
CC   direct and rapid detection in a screening process.
SQ   Sequence   114 BP;    0 A;    2 C;    2 G;    2 T;

  Query Match            2.8%;  Score 36;  DB 12;  Length 114;
  Best Local Similarity  4.5%;  Pred. No. 1.26e-06;
  Matches     5;  Conservative   33;  Mismatches  74;  Indels   0;  Gaps   0;

Db      3 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbtgcnnbnnnnnnnn 62
          :  :  :  :  :  :  :  :  :  :  :     :  :  :  :  :  : ||  :
Qy    933 TCTTCTGGGCTTCAGCTACCGTCTGGCCCTACACATCTACCAGCGTGGCCTGACTGGCTT 992

Db     63 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbtgcnnb 114
          :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :|||  :
Qy    993 CCTGGGCCAGGTGACCCGCTTTGTGGTGGACTTCATGCTGCATCACTGCATT 1044


Search completed: Thu Aug 21 13:23:47 1997
Job time : 416 secs.
```

****



****

Release 2.1D John F. Collins, Biocomputing Research Unit.

MPsrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Thu Aug 21 13:24:14 1997; MasPar time 668.58 Seconds
988.125 Million cell updates/sec

Tabular output not generated.

Title: >US-08-320-157-8
Description: (1-1287) from US08320157.seq
Perfect Score: 1287
N.A. Sequence: 1 TTTTAATATAAATTAATGTG.....CCTCAAGAGTACAGAAGCTT 1287
Comp: AAAATTATATTTAATTACAC.....GGAGTTCTCATGTCTTCGAA

Scoring table: TABLE default
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 707517 seqs, 256659390 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: EST-STS
1:EST1 2:EST2 3:EST3 4:EST4 5:EST5 6:EST6 7:EST7 8:EST8
9:EST9 10:EST10 11:EST11 12:EST12 13:EST13 14:EST14
15:EST15 16:EST16 17:EST17 18:EST18 19:EST19 20:EST20
21:EST21 22:EST22 23:EST23 24:EST24 25:EST25 26:EST26
27:EST27 28:EST28 29:EST29 30:EST30 31:EST31 32:EST32
33:EST33 34:EST34 35:EST35 36:EST36 37:EST37 38:EST38
39:EST39 40:EST40 41:EST41 42:EST42 43:EST43 44:EST44
45:EST45 46:EST46 47:EST47 48:EST48 49:EST49 50:EST50
51:EST51 52:EST52 53:EST53 54:EST54 55:EST55 56:EST56
57:EST57 58:EST58 59:EST59 60:EST60 61:EST61 62:EST62
63:EST63 64:EST64 65:EST65 66:EST66 67:EST67 68:EST68
69:EST69 70:EST70 71:EST71 72:EST72 73:EST73 74:EST74
75:EST75 76:EST76 77:EST77 78:EST78 79:EST79 80:EST80
81:EST81 82:EST82 83:EST83 84:EST84 85:EST85 86:EST86
87:EST87 88:EST88 89:EST89 90:EST90 91:EST91 92:EST92
93:EST93 94:EST94 95:EST95 96:EST96 97:EST97 98:EST98
99:EST99
Database: EST-STS-TWO
100:EST100 101:EST101 102:EST102 103:EST103 104:EST104
105:EST105 106:EST106 107:EST107 108:EST108 109:EST109
110:EST110 111:EST111 112:EST112 113:EST113 114:EST114
115:EST115 116:EST116 117:EST117 118:EST118 119:EST119
120:EST120 121:EST121 122:EST122 123:EST123 124:EST124
125:EST125 126:EST126 127:EST127 128:EST128 129:EST129
130:EST130 131:EST131 132:EST132 133:EST133 134:EST134
135:EST135 136:EST136 137:EST137 138:EST138 139:EST139
140:EST140 141:EST141 142:EST142 143:EST143 144:EST144
145:EST145 146:EST146 147:EST147 148:EST148 149:EST149
150:EST150 151:EST151 152:EST152 153:EST153 154:EST154
155:EST155 156:EST156 157:EST157 158:EST158 159:EST159
160:EST160 161:EST161 162:EST162 163:EST163 164:EST164
165:EST165 166:EST166 167:EST167 168:EST168 169:EST169
170:EST170 171:EST171 172:EST172 173:EST173 174:EST174
175:EST175 176:EST176 177:EST177 178:EST178 179:EST179
180:EST180 181:EST181 182:EST182 183:EST183 184:EST184
185:EST185 186:EST186 187:EST187 188:EST188 189:EST189
190:EST190 191:EST191 192:EST192 193:EST193 194:EST194
195:EST195 196:EST196 197:EST197 198:EST198

Statistics: Mean 11.600; Variance 2.183; scale 5.313

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description	Pred. No.
c	1	273	21.2	497	155	AA016399	mg88g02.r1 Soares mou	0.00e+00
	2	272	21.1	303	85	H56462	yt87a12.r1 Homo sapie	0.00e+00
c	3	254	19.7	545	153	W42014	mb16g04.r1 Soares mou	0.00e+00
c	4	232	18.0	446	138	AA049970	mj39b01.r1 Soares mou	0.00e+00
	5	72	5.6	292	138	AA049828	mj09a09.r1 Soares mou	4.27e-87
	6	72	5.6	361	139	AA050569	mj16b12.r1 Soares mou	4.27e-87
	7	60	4.7	232	133	N83998	KK4271F Homo sapiens	3.44e-64
	8	46	3.6	270	175	W82028	mf02e01.r1 Soares mou	7.53e-39
	9	39	3.0	320	186	AA139013	mr04a06.r1 Soares mou	5.11e-27
	10	31	2.4	279	82	H31840	EST106323 Rattus sp.	1.48e-14
	11	24	1.9	266	129	HSC07H082	H. sapiens partial cD	4.59e-05
c	12	24	1.9	337	109	HSC3LC012	H. sapiens partial cD	4.59e-05
c	13	24	1.9	358	133	N92010	za21a11.r1 Homo sapie	4.59e-05
c	14	24	1.9	395	54	H03526	yj37c07.r1 Homo sapie	4.59e-05
c	15	24	1.9	409	175	W84628	zd91f06.s1 Soares fet	4.59e-05
	16	25	1.9	411	44	H30391	ym58f11.s1 Homo sapie	2.60e-06
c	17	24	1.9	425	147	AA004647	zh92c08.s1 Soares fet	4.59e-05
	18	24	1.9	430	41	R25797	yg54b04.r1 Homo sapie	4.59e-05
	19	24	1.9	432	53	R92375	yq06g09.r1 Homo sapie	4.59e-05
c	20	24	1.9	435	117	W34930	mc60g11.r1 Soares mou	4.59e-05
c	21	24	1.9	446	170	W57984	zd18h07.s1 Soares fet	4.59e-05
	22	24	1.9	467	95	N41038	yy53d07.s1 Homo sapie	4.59e-05
	23	24	1.9	468	147	AA004387	zh92d01.r1 Soares fet	4.59e-05
	24	24	1.9	476	162	AA101808	zn44d07.s1 Stratagene	4.59e-05
c	25	24	1.9	478	110	HUM403H04B	Human fetal brain cDN	4.59e-05
c	26	24	1.9	480	136	AA001756	zh84b09.s1 Soares fet	4.59e-05
c	27	24	1.9	487	125	W74617	zd77e01.s1 Soares fet	4.59e-05
c	28	25	1.9	500	61	H15311	ym28b07.r1 Homo sapie	2.60e-06
c	29	23	1.8	240	48	HUM229F11B	Human aorta cDNA 5'-e	7.34e-04
	30	23	1.8	251	48	HUM250H12B	Human aorta cDNA 5'-e	7.34e-04
c	31	23	1.8	300	129	HSC0ME041	H. sapiens partial cD	7.34e-04
	32	23	1.8	301	186	AA138853	mr03a06.r1 Soares mou	7.34e-04
	33	23	1.8	317	78	R97647	yq59c09.s1 Homo sapie	7.34e-04
	34	23	1.8	317	29	R67853	yi28d10.s1 Homo sapie	7.34e-04
c	35	23	1.8	340	107	HSC2SG051	H. sapiens partial cD	7.34e-04
	36	23	1.8	368	194	AA169681	zo92f06.r1 Stratagene	7.34e-04
c	37	23	1.8	372	18	T50771	yb88e05.s1 Homo sapie	7.34e-04
c	38	23	1.8	377	19	T54824	yb43f08.r1 Homo sapie	7.34e-04
c	39	23	1.8	378	197	AA173916	zp01b08.s1 Stratagene	7.34e-04
	40	23	1.8	391	103	N72841	za03b02.s1 Homo sapie	7.34e-04
c	41	23	1.8	391	115	W17230	zb14a04.r1 Soares fet	7.34e-04
c	42	23	1.8	412	159	AA036957	zk32a09.s1 Soares pre	7.34e-04
c	43	23	1.8	445	174	W80365	zd82a03.s1 Soares fet	7.34e-04
c	44	23	1.8	490	138	AA046492	zk72c06.s1 Soares pre	7.34e-04
c	45	23	1.8	523	161	AA054723	zk68d08.s1 Soares pre	7.34e-04

ALIGNMENTS

RESULT 1
LOCUS AA016399 497 bp mRNA EST 02-AUG-1996
DEFINITION mg88g02.r1 Soares mouse embryo NbME13.5 14.5 Mus musculus cDNA clone 440114 5' similar to SW:BCLX_HUMAN Q07817 APOPTOSIS REGULATOR BCL-X. ;.
ACCESSION AA016399
NID g1478767

```
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae;
            Murinae; Mus.
REFERENCE   1  (bases 1 to 497)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT
            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:265450
            Possible reversed clone: similarity on wrong strand
            Seq primer: -28M13 rev2 from Amersham
            High quality sequence stop: 484.
FEATURES              Location/Qualifiers
     source           1..497
                      /organism="Mus musculus"
                      /strain="C57BL/6J"
                      /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                      polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                      was primed with a Not I - oligo(dT) primer [5'
                      TGTTACCAATCTGAAGTGGGAGCGGCCGCGGAAATTTTTTTTTTTTTTTTTTTTTT
                      T 3'], on equal amounts of mRNA from 2 13.5dpc and 2
                      14.5dpc embryos [total RNA provided by  Minoru Ko, Wayne
                      State Univ., from 2 ]; double-stranded cDNA was ligated to
                      Eco RI adaptors (Pharmacia), digested with Not I and
                      cloned into the Not I and Eco RI sites of the modified
                      pT7T3 vector. Library went through one round of
                      normalization, and was constructed by Bento Soares and
                      M.Fatima Bonaldo."
                      /clone="440114"
                      /clone_lib="Soares mouse embryo NbME13.5 14.5"
                      /sex="unknown"
                      /tissue_type="embryo"
                      /dev_stage="13.5-14.5dpc total fetus"
                      /lab_host="DH10B"
     mRNA             <1..>497
BASE COUNT      126 a    137 c    137 g     97 t
ORIGIN

  Query Match           21.2%;  Score 273;  DB 155;  Length 497;
  Best Local Similarity  78.7%;  Pred. No. 0.00e+00;
  Matches   374;  Conservative    0;  Mismatches 101;  Indels   0;  Gaps   0;

Db      22 tttagtccagccaggcacgtggctgggtgccccagatcttcacccgggagtcatgatctg 81
           || ||||||| ||||    ||    ||  ||||| |    | |  ||||||||||||| ||
Cp    1229 TTAAGTCCAGGCAGGGGTCTGAACTGGGACCCCAAAGGCACCCTTGGGAGTCATGATTTG 1170

Db      82 aagaatctgtgtaccacgaattggcccaacagaaccacaccaaaaatcaccattacggtc 141
           ||||||||  |||||||| || ||||||||||||||||||| | || |||||  ||| ||
Cp    1169 AAGAATCTTCGTACCACAAACTGGCCCAACAGAACCACACCCAGAACCACCAGCACGTTC 1110

Db     142 aggatggggtctctacgaaaattcagggctgccacccaaccgcctctctgtgcgatccat 201
           ||||||||  |  | |  || |||||||||||||||||| || || ||||||||  |||||
Cp    1109 AGGATGGGACCATTGCCCAAGTTCAGGGCTGCCACCCAGCCACCCCTCTGTGCAATCCAC 1050

Db     202 ctggcgatgtaatgatgcagtatgatatcagccaaaaagcaggtcacctggcccaggaag 261
           | ||| ||| | |||||||| ||||  ||  ||| ||||| ||||||||||||||||||
Cp    1049 CGGGCAATGCAGTGATGCAGCATGAAGTCCACCACAAAGCGGGTCACCTGGCCCAGGAAG 990

Db     262 ccggtcaaaccacgctggtagacgtacagggccagacggtagccaaagcccaggagagcc 321
           || ||||  |||||||||||||| ||  ||||||||||||||||  |||||||| ||||||
Cp     989 CCAGTCAGGCCACGCTGGTAGATGTGTAGGGCCAGACGGTAGCTGAAGCCCAGAAGAGCC 930

Db     322 accacgcggccccagctgatgccactcttaaataggctggaggcgatcttggtgaagagt 381
           ||||| ||||||||  |||||||||||| ||| |||||||||| ||||||||||||
Cp     929 ACCACACGGCCCCAATTGATGCCACTCTCAAACAGGCTGGAGGCAATCTTGGTGAAGTAC 870

Db     382 tcgtaggcattcccggctgtgggctgaagctgttctagtaaattctggaactctgtgtcg 441
           || |||||||||| | || |||||||| || || |  |  |   |||||||||||| |||
Cp     869 TCATAGGCATTCTCTGCCGTGGGCTGCAGGTGCTGCAACATGGTCTGGAACTCTGAGTCA 810

Db     442 tagcgccggttaatatcatctccgatgagagcaagctgccgacccacctgactca 496
           ||||| ||||| || || || || |||| || |||||||| |||||||| | ||
Cp     809 TAGCGTCGGTTGATGTCGTCCCCAATGATGGCGAGCTGCCGTCCCACCTGCCCCA 755


RESULT    2
LOCUS       H56462        303 bp    mRNA            EST       02-OCT-1995
DEFINITION  yt87a12.r1 Homo sapiens cDNA clone 231262 5'.
ACCESSION   H56462
NID         g1005106
KEYWORDS    EST.
SOURCE       clone=231262 primer=M13RP1 library=Soares pineal gland N3HPG
            vector=pT7T3D (Pharmacia) with a modified polylinker host=DH10B
            (ampicillin resistant) Rsite1=Not I Rsite2=Eco RI 1st strand cDNA
            was primed with a Not I - oligo(dT) primer [5'
            TGTTACCAATCTGAAGTGGGAGCGGCCGCCGTTTTTTTTTTTTTTTTTT 3'],
            double-stranded cDNA was size selected, ligated to Eco RI adapters
            (Pharmacia), digested with Not I and cloned into the Not I and Eco
            RI sites of a modified pT7T3 vector (Pharmacia). Library
            constructed by Bento Soares and M.Fatima Bonaldo.
  ORGANISM  Homo sapiens
            Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata;
            Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
            Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria;
            Eutheria; Archonta; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 303)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT
            Contact: Wilson RK
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            High quality sequence stops: 251
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
FEATURES              Location/Qualifiers
     source           1..303
                      /organism="Homo sapiens"
                      /clone="231262"
     mRNA             <1..>303
BASE COUNT       68 a     91 c     90 g     50 t      4 others
ORIGIN

  Query Match           21.1%;  Score 272;  DB 85;  Length 303;
  Best Local Similarity  95.7%;  Pred. No. 0.00e+00;
  Matches   291;  Conservative    0;  Mismatches  11;  Indels   2;  Gaps   2;

Db       1 gggcaaggcaatgnnncctcccaggcaggagtgcggagagcctgccctgccctctgcttct 60
           |||||||||   | ||||||||||||||||||||||||||||||||||||||||||||||
Qy     553 GGGCAAGGCCCAGGTCCTCCCAGGCAGGAGTGCGGAGAGCCTGCCCTGCCCTCTGCTTCT 612
```

```
Db     61 gaggagcaggtagcccaggacacagaggaggttttccgcagctacgttttttaccaccat 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    613 GAGGAGCAGGTAGCCCAGGACACAGAGGAGGTTTTCCGCAGCTACGTTTTTTACCACCAT 672

Db    121 cagcaggaacaggaggctgaaggggcggctgctgacnc-gacccagagatggtcaccttta 179
          |||||||||||||||||||||||||||||||||||     | ||||||||||||||||||||
Qy    673 CAGCAGGAACAGGAGGCTGAAGGGGCGGCTGCCCCTGCCGACCCAGAGATGGTCACCTTA 732

Db    180 cctctgcaacctagcagcaccatggggcaggtgggacggcagctcgccatcattggggac 239
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    733 CCTCTGCAACCTAGCAGCACCATGGGGCAGGTGGGACGGCAGCTCGCCATCATTGGGGAC 792

Db    240 gacatcaaccgacgctatgactcagagttccagaccatgttncagcacctgcagtcccac 299
          |||||||||||||||||||||||||||||||||||||||| |||||||||||| |||||
Qy    793 GACATCAACCGACGCTATGACTCAGAGTTCCAGACCATGTTGCAGCACCTGCAG-CCCAC 851

Db    300 ggca 303
          ||||
Qy    852 GGCA 855
```

```
RESULT    3
LOCUS       W42014         545 bp    mRNA            EST       11-SEP-1996
DEFINITION  mb16g04.r1 Soares mouse p3NMF19.5 Mus musculus cDNA clone 329622 5'
            similar to SW:BCLX_HUMAN Q07817 APOPTOSIS REGULATOR BCL-X. ;.
ACCESSION   W42014
NID         g1326554
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Mus.
REFERENCE   1  (bases 1 to 545)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT
            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:211022
            Possible reversed clone: similarity on wrong strand
            Seq primer: ETPrimer
            High quality sequence stop: 411.
FEATURES              Location/Qualifiers
     source           1..545
                      /organism="Mus musculus"
                      /note="Vector: pT7T3D (Pharmacia) with a modified
                      polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                      was primed with a Not I - oligo(dT) primer [5'
                      TGTTACCAATCTGAAGTGGGAGCGGCCGCATTTTTTTTTTTTTTTTTT 3'],
                      double-stranded cDNA was size selected, ligated to Eco RI
                      adapters (Pharmacia), digested with Not I and cloned into
                      the Not I and Eco RI sites of a modified pT7T3 vector
                      (Pharmacia). Library went through one round of
                      normalization to a Cot = 5. Library constructed by Bento
                      Soares and M.Fatima Bonaldo. RNA was kindly provided by
                      Dr. Minoru Ko (Wayne State University)."
                      /clone="329622"
                      /clone_lib="Soares mouse p3NMF19.5"
                      /dev_stage="19.5 dpc total fetus"
                      /lab_host="DH10B (ampicillin resistant)"
     mRNA             <1..>545
BASE COUNT      134 a    139 c    165 g    105 t      2 others
ORIGIN
```

```
  Query Match           19.7%;  Score 254;  DB 153;  Length 545;
  Best Local Similarity  78.5%;  Pred. No. 0.00e+00;
  Matches   384;  Conservative    0;  Mismatches 100;  Indels   5;  Gaps   5;

Db     31 tttagtccagccaggcacgtgg-ctggggccccagtgggtcccc-gggagtcatgatctg 88
          || ||||||| ||||    ||  ||||| |||||  ||  |||  ||||||||||| ||
Cp   1229 TTAAGTCCAGGCAGGGGTCTGAACTGGGACCCCAAAGGCACCCTTGGGAGTCATGATTTG 1170

Db     89 aagaatctgtgtaccacgaattggcccaacagaaccacaccaaaaatcaccattacggtc 148
          ||||||||  ||||||| || ||||||||||||||||||| | || |||||  ||| ||
Cp   1169 AAGAATCTTCGTACCACAAACTGGCCCAACAGAACCACACCCAGAACCACCAGCACGTTC 1110

Db    149 aggatggggactctacgaaaattcagggctgccacccaaccgcctctctgtgcgatccat 208
          ||||||||  |  | |  || |||||||||||||||||| || || |||||||| |||||
Cp   1109 AGGATGGGACCATTGCCCAAGTTCAGGGCTGCCACCCAGCCACCCCTCTGTGCAATCCAC 1050

Db    209 ctggcgatgtaatgatgcagtatgatatcagccaaaaagcaggtcacctggcccaggaag 268
          | ||| ||| | |||||||| ||||  ||  ||| ||||| ||||||||||||||||||
Cp   1049 CGGGCAATGCAGTGATGCAGCATGAAGTCCACCACAAAGCGGGTCACCTGGCCCAGGAAG 990

Db    269 ccggtcaaaccacgctggtagacgtacagggccagacggtagccaaagcccaggagagcc 328
          || |||| |||||||||||||| ||  |||||||||||||||  |||||||| ||||||
Cp    989 CCAGTCAGGCCACGCTGGTAGATGTGTAGGGCCAGACGGTAGCTGAAGCCCAGAAGAGCC 930

Db    329 accacgcggccccagctgatgccactcttaaataggctggaggcgatcttggtgaagagt 388
          ||||| ||||||||  |||||||||||| ||| ||||||||||| ||||||||||
Cp    929 ACCACACGGCCCCAATTGATGCCACTCTCAAACAGGCTGGAGGCAATCTTGGTGAAGTAC 870

Db    389 tcgtaggcattcccggctgtgggctgaagctgttctagtagattctggaactctgtgtcg 448
          || |||||||| | || |||||||| || || |  |  |   |||||||||||| |||
Cp    869 TCATAGGCATTCTCTGCCGTGGGCTGCAGGTGCTGCAACATGGTCTGGAACTCTGAGTCA 810

Db    449 ta-cgccggttaatatca-ctccgatgagagcaagctgccgacc-acgttgaccaagatg 505
          || || ||||| || ||  | || ||||  || |||||||| || || |   ||| | ||
Cp    809 TAGCGTCGGTTGATGTCGTCCCCAATGATGGCGAGCTGCCGTCCCACCTGCCCCATGGTG 750

Db    506 ctgttgggt 514
          ||| | |||
Cp    749 CTGCTAGGT 741
```

```
RESULT    4
LOCUS       AA049970       446 bp    mRNA            EST       09-SEP-1996
DEFINITION  mj39b01.r1 Soares mouse embryo NbME13.5 14.5 Mus musculus cDNA
            clone 478441 5' similar to SW:BCLX_HUMAN Q07817 APOPTOSIS REGULATOR
            BCL-X. ;.
ACCESSION   AA049970
NID         g1529641
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Mus.
REFERENCE   1  (bases 1 to 446)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT
            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
```

```
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:289185
            Possible reversed clone: similarity on wrong strand
            Seq primer: -28M13 rev2 from Amersham
            High quality sequence stop: 319.
FEATURES              Location/Qualifiers
     source           1..446
                      /organism="Mus musculus"
                      /strain="C57BL/6J"
                      /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                      polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                      was primed with a Not I - oligo(dT) primer [5'
                      TGTTACCAATCTGAAGTGGGAGCGGCCGCGGAAATTTTTTTTTTTTTTTTTTTTTT
                      T 3'], on equal amounts of mRNA from 2 13.5dpc and 2
                      14.5dpc embryos [total RNA provided by  Minoru Ko, Wayne
                      State Univ., from 2 ]; double-stranded cDNA was ligated to
                      Eco RI adaptors (Pharmacia), digested with Not I and
                      cloned into the Not I and Eco RI sites of the modified
                      pT7T3 vector. Library went through one round of
                      normalization, and was constructed by Bento Soares and
                      M.Fatima Bonaldo."
                      /clone="478441"
                      /clone_lib="Soares mouse embryo NbME13.5 14.5"
                      /sex="unknown"
                      /tissue_type="embryo"
                      /dev_stage="13.5-14.5dpc total fetus"
                      /lab_host="DH10B"
     mRNA             <1..>446
BASE COUNT      110 a    118 c    127 g     91 t
ORIGIN

  Query Match           18.0%;  Score 232;  DB 138;  Length 446;
  Best Local Similarity  80.9%;  Pred. No. 0.00e+00;
  Matches   304;  Conservative    0;  Mismatches  72;  Indels   0;  Gaps   0;

Db     71 gtcatgatctgaagaatctgtgtaccacgaattggcccaacagaaccacaccaaaaatca 130
        |||||||| ||||||||||  ||||||| || ||||||||||||||||||||| | || ||
Cp   1180 GTCATGATTTGAAGAATCTTCGTACCACAAACTGGCCCAACAGAACCACACCCAGAACCA 1121

Db    131 ccattacggtcatgatggggtctctacgaaaattcagggctgccacccaaccgcctctct 190
          |||  ||| ||| ||||||  |  | |  || ||||||||||||||||| || || ||||
Cp   1120 CCAGCACGTTCAGGATGGGACCATTGCCCAAGTTCAGGGCTGCCACCCAGCCACCCCTCT 1061

Db    191 gtgcgatccatctggcgatgtaatgatgcagtatgatatcagccaaaaagcaggtcacct 250
          |||| ||||| | ||| ||| | |||||||| ||||  ||  ||| ||||| ||||||||
Cp   1060 GTGCAATCCACCGGGCAATGCAGTGATGCAGCATGAAGTCCACCACAAAGCGGGTCACCT 1001

Db    251 ggcccaggaagccggtcaaaccacgctggtagacgtacagggccagacggtagccaaagc 310
          |||||||||||||| ||||  |||||||||||||| ||  ||||||||||||||||  ||||
Cp   1000 GGCCCAGGAAGCCAGTCAGGCCACGCTGGTAGATGTGTAGGGCCAGACGGTAGCTGAAGC 941

Db    311 ccaggagagccaccacgcggccccagctgatgccactcttaaataggctggaggcgatct 370
          |||| |||||||||||| ||||||||  |||||||||||| ||| ||||||||||| ||||
Cp    940 CCAGAAGAGCCACCACACGGCCCCAATTGATGCCACTCTCAAACAGGCTGGAGGCAATCT 881

Db    371 tggtgaagagttcgtaggcatttccggctgtgggctgaagctgttctagtaaattcttga 430
          |||||||   || ||||||||  | || |||||||| || || |  |  |  ||| ||
Cp    880 TGGTGAAGTACTCATAGGCATTCTCTGCCGTGGGCTGCAGGTGCTGCAACATGGTCTGGA 821

Db    431 actctgtgtcgtagcg 446
          |||||| ||| |||||
Cp    820 ACTCTGAGTCATAGCG 805


RESULT    5
LOCUS       AA049828      292 bp    mRNA            EST       09-SEP-1996
DEFINITION  mj09a09.r1 Soares mouse embryo NbME13.5 14.5 Mus musculus cDNA
            clone 475576 5'.
ACCESSION   AA049828
NID         g1529499
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Mus.
REFERENCE   1  (bases 1 to 292)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT
            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:286320
            Seq primer: -28M13 rev2 from Amersham.
FEATURES              Location/Qualifiers
     source           1..292
                      /organism="Mus musculus"
                      /strain="C57BL/6J"
                      /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                      polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                      was primed with a Not I - oligo(dT) primer [5'
                      TGTTACCAATCTGAAGTGGGAGCGGCCGCGGAAATTTTTTTTTTTTTTTTTTTTTT
                      T 3'], on equal amounts of mRNA from 2 13.5dpc and 2
                      14.5dpc embryos [total RNA provided by  Minoru Ko, Wayne
                      State Univ., from 2 ]; double-stranded cDNA was ligated to
                      Eco RI adaptors (Pharmacia), digested with Not I and
                      cloned into the Not I and Eco RI sites of the modified
                      pT7T3 vector. Library went through one round of
                      normalization, and was constructed by Bento Soares and
                      M.Fatima Bonaldo."
                      /clone="475576"
                      /clone_lib="Soares mouse embryo NbME13.5 14.5"
                      /sex="unknown"
                      /tissue_type="embryo"
                      /dev_stage="13.5-14.5dpc total fetus"
                      /lab_host="DH10B"
     mRNA             <1..>292
BASE COUNT       67 a     88 c     83 g     54 t
ORIGIN

  Query Match           5.6%;  Score 72;  DB 138;  Length 292;
  Best Local Similarity  79.7%;  Pred. No. 4.27e-87;
  Matches   137;  Conservative    0;  Mismatches  29;  Indels   6;  Gaps   4;

Db    127 ctccaccaagacctgaaaaatggcatctggacaaggaccaggtcccccgaaggtgggctg 186
          ||||||  |||||||||||||||| || || ||||| |||||||| || | |  ||  ||
Qy    525 CTCCACTGAGACCTGAAAAATGGCTTCGGGGCAAGGCCCAGGTCCTCCCAGGCAGGAGTG 584

Db    187 cgatgagtc--ccc-gtcc-c--cttctgaacagcaggttgcccaggacacagaggaggt 240
          ||  ||| |  ||| | || |  |||||||  ||||||| ||||||||||||||||||||
Qy    585 CGGAGAGCCTGCCCTGCCCTCTGCTTCTGAGGAGCAGGTAGCCCAGGACACAGAGGAGGT 644

Db    241 ctttcgaagctacgtttttttacctccaccagcaggaacaggagacccagggg 292
          || || |||||||||||||||| ||| |||||||||||||| |  | |||
Qy    645 TTTCCGCAGCTACGTTTTTTTACCACCATCAGCAGGAACAGGAGGCTGAAGGG 696


RESULT    6
LOCUS       AA050569      361 bp    mRNA            EST       09-SEP-1996
DEFINITION  mj16b12.r1 Soares mouse embryo NbME13.5 14.5 Mus musculus cDNA
            clone 476255 5'.
```

```
ACCESSION   AA050569
NID         g1530241
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Mus.
REFERENCE   1  (bases 1 to 361)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT
            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:286999
            Seq primer: -28M13 rev2 from Amersham.
FEATURES             Location/Qualifiers
     source          1..361
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGGAAATTTTTTTTTTTTTTTTTTTTTT
                     T 3'], on equal amounts of mRNA from 2 13.5dpc and 2
                     14.5dpc embryos [total RNA provided by  Minoru Ko, Wayne
                     State Univ., from 2 ]; double-stranded cDNA was ligated to
                     Eco RI adaptors (Pharmacia), digested with Not I and
                     cloned into the Not I and Eco RI sites of the modified
                     pT7T3 vector. Library went through one round of
                     normalization, and was constructed by Bento Soares and
                     M.Fatima Bonaldo."
                     /clone="476255"
                     /clone_lib="Soares mouse embryo NbME13.5 14.5"
                     /sex="unknown"
                     /tissue_type="embryo"
                     /dev_stage="13.5-14.5dpc total fetus"
                     /lab_host="DH10B"
     mRNA            <1..>361
BASE COUNT       87 a    108 c    103 g     63 t
ORIGIN

  Query Match             5.6%;  Score 72;  DB 139;  Length 361;
  Best Local Similarity  79.7%;  Pred. No. 4.27e-87;
  Matches   137;  Conservative    0;  Mismatches  29;  Indels   6;  Gaps   4;

Db    196 ctccaccaagacctgaaaaatggcatctggacaaggaccaggtcccccgaaggtgggctg 255
          ||||||  |||||||||||||||| || || ||||| |||||||| || | |  ||  ||
Qy    525 CTCCACTGAGACCTGAAAAATGGCTTCGGGGCAAGGCCCAGGTCCTCCCAGGCAGGAGTG 584

Db    256 cgatgagtc--ccc-gtcc-c--cttctgaacagcaggttgcccaggacacagaggaggt 309
          ||  ||| |  ||| | || |  |||||||  ||||||| ||||||||||||||||||||
Qy    585 CGGAGAGCCTGCCCTGCCCTCTGCTTCTGAGGAGCAGGTAGCCCAGGACACAGAGGAGGT 644

Db    310 ctttcgaagctacgtttttttacctccaccagcaggaacaggagacccagggg 361
           || || |||||||||||||| ||| ||||||||||||| |  | |||
Qy    645 TTTCCGCAGCTACGTTTTTTACCACCATCAGCAGGAACAGGAGGCTGAAGGG 696


RESULT    7
LOCUS       N83998        232 bp    mRNA            EST       01-APR-1996
DEFINITION  KK4271F Homo sapiens cDNA clone KK4271 5' similar to BAK FOR BCL-2
            HOMOLOGUE.
ACCESSION   N83998
NID         g1259623
KEYWORDS    EST.
SOURCE      human clone=KK4271 primer=GAAATTAACCCTCACTAAAGGG library=Fetal
            heart, Lambda ZAP Express vector=Lambda ZAP Express host=E. coli
            XL1-Blue Rsite1=EcoRI Rsite2=XhoI mRNA was purified from human
            fetal hearts (8-10 weeks).  cDNA was synthesized using a XhoI-Oligo
            dT adaptor-primer.  EcoRI adaptors were ligated, followed by
            digestion with XhoI, for directional cloning into predigested
            lambda ZAP Express.
  ORGANISM  Homo sapiens
            Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata;
            Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
            Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria;
            Eutheria; Archonta; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 232)
  AUTHORS   Liew,C.C.
  TITLE     cDNAs from fetal heart
  JOURNAL   Unpublished (1996)
COMMENT
            Contact: Liew CC
            Molecular Cardiology
            University of Toronto
            Banting Institute, 100 College St., Toronto, Ontario, M5G1L5
            Tel: 4169788758
            Fax: 4169785650
            Email: liewcc@utcc.utoronto.ca.
FEATURES             Location/Qualifiers
     source          1..232
                     /organism="Homo sapiens"
                     /clone="KK4271"
     mRNA            <1..>232
BASE COUNT       52 a     69 c     57 g     54 t
ORIGIN

  Query Match             4.7%;  Score 60;  DB 133;  Length 232;
  Best Local Similarity  96.9%;  Pred. No. 3.44e-64;
  Matches    62;  Conservative    0;  Mismatches   2;  Indels   0;  Gaps   0;

Db     87 tcagaggagcaggtagcccaggacacagaggaggttttccgcagctacgttttttaccgc 146
          || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Qy    610 TCTGAGGAGCAGGTAGCCCAGGACACAGAGGAGGTTTTCCGCAGCTACGTTTTTTACCAC 669

Db    147 catc 150
          ||||
Qy    670 CATC 673


RESULT    8
LOCUS       W82028        270 bp    mRNA            EST       12-SEP-1996
DEFINITION  mf02e01.r1 Soares mouse p3NMF19.5 Mus musculus cDNA clone 403896
            5'.
ACCESSION   W82028
NID         g1393105
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Mus.
REFERENCE   1  (bases 1 to 270)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT
            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
```

```
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:247664
            Seq primer: -28M13 rev2 from Amersham.
FEATURES              Location/Qualifiers
     source           1..270
                      /organism="Mus musculus"
                      /note="Vector: pT7T3D (Pharmacia) with a modified
                      polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                      was primed with a Not I - oligo(dT) primer [5'
                      TGTTACCAATCTGAAGTGGGAGCGGCCGCATTTTTTTTTTTTTTTTTT 3'],
                      double-stranded cDNA was size selected, ligated to Eco RI
                      adapters (Pharmacia), digested with Not I and cloned into
                      the Not I and Eco RI sites of a modified pT7T3 vector
                      (Pharmacia). Library went through one round of
                      normalization to a Cot = 5. Library constructed by Bento
                      Soares and M.Fatima Bonaldo. RNA was kindly provided by
                      Dr. Minoru Ko (Wayne State University)."
                      /clone="403896"
                      /clone_lib="Soares mouse p3NMF19.5"
                      /dev_stage="19.5 dpc total fetus"
                      /lab_host="DH10B (ampicillin resistant)"
     mRNA             <1..>270
BASE COUNT       61 a     83 c     74 g     52 t
ORIGIN

  Query Match             3.6%;  Score 46;  DB 175;  Length 270;
  Best Local Similarity  76.7%;  Pred. No. 7.53e-39;
  Matches   102;  Conservative    0;  Mismatches  26;  Indels   5;  Gaps   4;

Db    143 ctccaccaagacctgaaaaatggcatctggacaaggaccaggtcccccgaaggtgg-gt- 200
          ||||||  |||||||||||||||| || || ||||| |||||||| || | |  || ||
Qy    525 CTCCACTGAGACCTGAAAAATGGCTTCGGGGCAAGGCCCAGGTCCTCCCAGGCAGGAGTG 584

Db    201 ctgcgatgagtccccgtcc-c--cttctgaacagcaggttgcccaggacacagaggaggt 257
          | | ||     ||| | || |  |||||||  |||||||  ||||||||||||||||||
Qy    585 CGGAGAGCCTGCCCTGCCCTCTGCTTCTGAGGAGCAGGTAGCCCAGGACACAGAGGAGGT 644

Db    258 ctttcgaagctac 270
            || || ||||||
Qy    645 TTTCCGCAGCTAC 657


RESULT    9
LOCUS       AA139013      320 bp    mRNA            EST       02-DEC-1996
DEFINITION  mr04a06.r1 Soares mouse 3NbMS Mus musculus cDNA clone 596434 5'
            similar to TR:G595926 G595926 BAK-2 PROTEIN. ;.
ACCESSION   AA139013
NID         g1701240
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Mus.
REFERENCE   1  (bases 1 to 320)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT
            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:361866
            Seq primer: -28M13 rev2 from Amersham.
FEATURES              Location/Qualifiers
     source           1..320
                      /organism="Mus musculus"
                      /strain="C57BL/6J"
                      /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                      polylinker; Site_1: Not I; Site_2: Eco RI; [5'
                      TGTTACCAATCTGAAGTGGGAGCGGCCGCGCTGTTTTTTTTTTTTTTTTTT
                      3']; double-stranded cDNA was ligated to Eco RI adaptors
                      (Pharmacia), digested with Not I and cloned into the Not
                      I and Eco RI sites of the modified pT7T3 vector. RNA
                      provided by Dr. Bertrand Jordan. Library went through
                      three rounds of normalization, and was constructed by
                      Bento Soares and M.Fatima Bonaldo."
                      /clone="596434"
                      /clone_lib="Soares mouse 3NbMS"
                      /sex="male"
                      /dev_stage="4 weeks"
                      /lab_host="DH10B"
     mRNA             <1..>320
BASE COUNT       68 a     92 c     96 g     64 t
ORIGIN

  Query Match             3.0%;  Score 39;  DB 186;  Length 320;
  Best Local Similarity  84.2%;  Pred. No. 5.11e-27;
  Matches    48;  Conservative    0;  Mismatches   9;  Indels   0;  Gaps   0;

Db    264 ccgtccccttctgaacagcaggttgcccaggacacagaggaggtctttcgaagctac 320
          || ||  |||||||  ||||||| |||||||||||||||||||| || || ||||||
Qy    601 CCCTCTGCTTCTGAGGAGCAGGTAGCCCAGGACACAGAGGAGGTTTTCCGCAGCTAC 657


RESULT   10
LOCUS       H31840        279 bp    mRNA            EST       08-SEP-1995
DEFINITION  EST106323 Rattus sp. cDNA 5' end.
ACCESSION   H31840
NID         g977257
KEYWORDS    EST.
SOURCE      rat primer=M13 Reverse library=Rat PC-12 cells, untreated
            vector=pBluescript SK- Rsite1=EcoRI Rsite2=XhoI poly(A)+ RNA was
            purified from untreated PC12 cells cultured for 9 days.  cDNA was
            constructed using an oligo-dT primer and directionally cloned using
            the Lambda ZAP II Vector Kit by Stratagene.
  ORGANISM  Rattus sp.
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Myomorpha;
            Muridae; Murinae; Rattus.
REFERENCE   1  (bases 1 to 279)
  AUTHORS   Lee,N.H., Weinstock,K.G., Kirkness,E.F., Earle-Hughes,J.A.,
            Fuldner,R.A., Marmaras,S., Glodek,A., Gocayne,J.D., Adams,M.D.,
            Kerlavage,A.R., Fraser,C.M. and Venter,J.C.
  TITLE     Comparative expressed sequence tag analysis of differential gene
            expression profiles in PC-12 cells before and after nerve growth
            factor treatment
  JOURNAL   Proc. Natl. Acad. Sci. USA 92, 8303-8307 (1995)
COMMENT     Other_ESTs: EST106322
            Contact: Lee NH
            The Institute for Genomic Research
            932 Clopper Rd, Gaithersburg, MD 20878
            Tel: 3018699056
            Fax: 3018699423
            Email: nhlee@tigr.org
            For clone availability please contact  the TIGR Database
            (tdbinfo@tdb.tigr.org).
FEATURES              Location/Qualifiers
     source           1..279
                      /organism="Rattus sp."
```

```
     mRNA            <1..>279
BASE COUNT       65 a     80 c     76 g     53 t      5 others
ORIGIN

  Query Match             2.4%;  Score 31;  DB 82;  Length 279;
  Best Local Similarity  82.7%;  Pred. No. 1.48e-14;
  Matches    43;  Conservative    0;  Mismatches   8;  Indels   1;  Gaps   1;

Db    226 cctccaccaagaccggaaaaatggcatccgggacaaggaccaggtnctncca 277
        |||||||  ||||| |||||||||| || ||| ||||| |||||| || |||
Qy    524 CCTCCACTGAGACCTGAAAAATGGCTTCGGGG-CAAGGCCCAGGTCCTCCCA 574


RESULT   11
LOCUS       HSC07H082     266 bp    RNA             EST       21-SEP-1995
DEFINITION  H. sapiens partial cDNA sequence; clone c-07h08.
ACCESSION   F01610
NID         g645167
KEYWORDS    partial cDNA sequence; transcribed sequence fragment.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 266)
  AUTHORS   Genexpress.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-JAN-1995) Genethon, B.P. 60, 91002 Evry Cedex France
            and Genetique Moleculaire et Biologie du developpement, CNRS UPR420
            B.P. 8, 94801 Villejuif Cedex France.E-mail: genexpress@genethon.fr
REFERENCE   2  (bases 1 to 266)
  AUTHORS   Genexpress.
  TITLE     The Genexpress cDNA program
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 266)
  AUTHORS   Auffray,C., Behar,G., Bois,F., Bouchier,C., da Silva,C.,
            Devignes,M.D., Duprat,S., Houlgatte,R., Jumeau,M.N., Lamy,B.,
            Lorenzo,F., Mitchell,H., Mariage-Samson,R., Pietu,G., Pouliot,Y.,
            Sebastiani-Kabaktchis,C. and Tessier,A.
  TITLE     IMAGE: molecular integration of the analysis of the human genome
            and its expression
  JOURNAL   C. R. Acad. Sci. III, Sci. Vie 318 (2), 263-272 (1995)
  MEDLINE   95277534
COMMENT     Cloning_method: total mRNA was oligo-(dT) primed and directionally
            cloned 5' -> 3' into the HindIII -> NotI sites of the lafmid BA
            vector;
            Sequencing_method: single read, full automatic;
            Primer: (-21)M13_universal;
            cDNA sequence complementary to mRNA (3'end)
            Stretch_removed:  removed at sequence 5'end
            Normalization_method: Bento Soares, P.N.A.S in press;
            Genexpress_library_idt: C;
            Genexpress_sequence_idt: a1c-07h08.
FEATURES              Location/Qualifiers
     source           1..266
                      /organism="Homo sapiens"
                      /isolate="muscular atrophy patient"
                      /clone_lib="normalized infant brain cDNA from B.Soares,
                      Psychiatry Dept. Columbia University USA"
                      /sex="female"
                      /tissue_type="total brain"
                      /dev_stage="3 months old"
BASE COUNT       94 a     47 c     49 g     75 t      1 others
ORIGIN

  Query Match             1.9%;  Score 24;  DB 129;  Length 266;
  Best Local Similarity  80.0%;  Pred. No. 4.59e-05;
  Matches    32;  Conservative    0;  Mismatches   8;  Indels   0;  Gaps   0;

Db    161 aataaaaatttaaatatttttgaactgaataataataaaa 200
          |||||||||| |||| || | |||||||| || |  ||||
Qy     55 AATAAAAATTCAAATGTTATAGAACTGAAAAAGATGAAAA 94


RESULT   12
LOCUS       HSC3LC012     337 bp    RNA             EST       21-SEP-1995
DEFINITION  H. sapiens partial cDNA sequence; clone c-31c01.
ACCESSION   F10881
NID         g709398
KEYWORDS    partial cDNA sequence; transcribed sequence fragment.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 337)
  AUTHORS   Genexpress.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-JAN-1995) Genethon, B.P. 60, 91002 Evry Cedex France
            and Genetique Moleculaire et Biologie du developpement, CNRS UPR420
            B.P. 8, 94801 Villejuif Cedex France. E-mail:
            genexpress@genethon.fr
REFERENCE   2  (bases 1 to 337)
  AUTHORS   Genexpress.
  TITLE     The Genexpress cDNA program
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 337)
  AUTHORS   Auffray,C., Behar,G., Bois,F., Bouchier,C., da Silva,C.,
            Devignes,M.D., Duprat,S., Houlgatte,R., Jumeau,M.N., Lamy,B.,
            Lorenzo,F., Mitchell,H., Mariage-Samson,R., Pietu,G., Pouliot,Y.,
            Sebastiani-Kabaktchis,C. and Tessier,A.
  TITLE     IMAGE: molecular integration of the analysis of the human genome
            and its expression
  JOURNAL   C. R. Acad. Sci. III, Sci. Vie 318 (2), 263-272 (1995)
  MEDLINE   95277534
COMMENT     Cloning_method: total mRNA was oligo-(dT) primed and directionally
            cloned 5' -> 3' into the HindIII -> NotI sites of the lafmid BA
            vector;
            Sequencing_method: single read, full automatic;
            Primer: (-21)M13_universal;
            cDNA sequence complementary to mRNA (3'end)
            Stretch_removed:  removed at sequence 5'end
            Normalization_method: Bento Soares, P.N.A.S. 91:9228-9232(1994);
            Genexpress_library_idt: C;
            Genexpress_sequence_idt: a1c-31c01.
FEATURES              Location/Qualifiers
     source           1..337
                      /organism="Homo sapiens"
                      /isolate="muscular atrophy patient"
                      /clone_lib="normalized infant brain cDNA from B.Soares,
                      Psychiatry Dept. Columbia University USA"
                      /sex="female"
                      /tissue_type="total brain"
                      /dev_stage="3 months old"
BASE COUNT      101 a     78 c     92 g     64 t      2 others
ORIGIN

  Query Match             1.9%;  Score 24;  DB 109;  Length 337;
  Best Local Similarity  69.4%;  Pred. No. 4.59e-05;
  Matches    43;  Conservative    0;  Mismatches  19;  Indels   0;  Gaps   0;

Db    266 ggaaaacatcctgggtggtctgggctaaatgatcaagtgaaacagaaagccaggaaggat 325
          ||||||| ||||  |||  |||||||||  || ||    ||| ||||  ||  || ||| |
Cp    649 GGAAAACCTCCTCTGTGTCCTGGGCTACCTGCTCCTCAGAAGCAGAGGGCAGGGCAGGCT 590

Db    326 ct 327
          ||
Cp    589 CT 588


RESULT   13
LOCUS       N92010        358 bp    mRNA            EST       04-APR-1996
DEFINITION  za21a11.r1 Homo sapiens cDNA clone 293180 5' similar to
            gb:M10942_cds1 Human metallothionein-Ie gene (HUMAN);.
ACCESSION   N92010
NID         g1264319
KEYWORDS    EST.
SOURCE      human clone=293180 primer=mob.REGA+ET library=Soares fetal liver
```

```
            spleen 1NFLS vector=pT7T3D (Pharmacia) with a modified polylinker
            host=DH10B (ampicillin resistant) Rsite1=Pac I Rsite2=Eco RI Liver
            and spleen from a 20 week-post conception male fetus. 1st strand
            cDNA was primed with a Pac I - oligo(dT) primer [5'
            AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTTT 3'], double-stranded
            cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac
            I and cloned into the Pac I and Eco RI sites of the modified pT7T3
            vector.  Library went through one round of normalization. Library
            constructed by Bento Soares and M.Fatima Bonaldo.
  ORGANISM  Homo sapiens
            Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata;
            Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
            Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria;
            Eutheria; Archonta; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 358)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT
            Contact: Wilson RK
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
FEATURES              Location/Qualifiers
     source           1..358
                      /organism="Homo sapiens"
                      /clone="293180"
     mRNA             <1..>358
BASE COUNT       83 a     89 c     82 g     89 t     15 others
ORIGIN

  Query Match             1.9%;  Score 24;  DB 133;  Length 358;
  Best Local Similarity  75.6%;  Pred. No. 4.59e-05;
  Matches    34;  Conservative    0;  Mismatches  11;  Indels    0;  Gaps    0;

Db    126 gagctgctgttcctgctgccccgtgntgctgcaagtgtgcccagg 170
          ||||||| || ||  ||||||| || |||||| ||  |||  |||
Cp    777 GAGCTGCCGTCCCACCTGCCCCATGGTGCTGCTAGGTTGCAGAGG 733


RESULT   14
LOCUS       H03526         395 bp    mRNA            EST       20-JUN-1995
DEFINITION  yj37c07.r1 Homo sapiens cDNA clone 150924 5'.
ACCESSION   H03526
NID         g866459
KEYWORDS    EST.
SOURCE      human clone=150924 library=Soares placenta Nb2HP vector=pT7T3D
            (Pharmacia) with a modified polylinker host=DH10B (ampicillin
            resistant) primer=M13RP1 Rsite1=Not I Rsite2=Eco RI Female placenta
            obtained at birth (full term). 1st strand cDNA was primed with a
            Not I - oligo(dT) primer [5'
            AACTGGAAGAATTCGCGGCCGCAGGAATTTTTTTTTTTTTTTTT 3'], double-stranded
            cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not
            I and cloned into the Not I and Eco RI sites of the modified pT7T3
            vector. Library went through one round of normalization. Library
            constructed by Bento Soares and M.Fatima Bonaldo.  .
  ORGANISM  Homo sapiens
            Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata;
            Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
            Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria;
            Eutheria; Archonta; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 395)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT
            Contact: Wilson RK
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            High qality sequence stops: 314
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
FEATURES              Location/Qualifiers
     source           1..395
                      /organism="Homo sapiens"
                      /clone="150924"
BASE COUNT       81 a     95 c     79 g    139 t      1 others
ORIGIN

  Query Match             1.9%;  Score 24;  DB 54;  Length 395;
  Best Local Similarity  80.0%;  Pred. No. 4.59e-05;
  Matches    32;  Conservative    0;  Mismatches   8;  Indels    0;  Gaps    0;

Db    333 gttgtttttttctttctttcttttttctcttctataatttttt 372
          ||||||||| ||||   ||||||||  ||||||||  |||
Cp    105 GTTGTTTTTACTTTTCATCTTTTTCAGTTCTATAACATTT 66


RESULT   15
LOCUS       W84628         409 bp    mRNA            EST       17-OCT-1996
DEFINITION  zd91f06.s1 Soares fetal heart NbHH19W Homo sapiens cDNA clone
            356867 3'.
ACCESSION   W84628
NID         g1395739
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 409)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT
            Contact: Wilson RK
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 1099   Std Error: 0.00
            Seq primer: mob.REGA+ET.
FEATURES              Location/Qualifiers
     source           1..409
                      /organism="Homo sapiens"
                      /note="Organ: heart; Vector: pT7T3D (Pharmacia) with a
                      modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                      strand cDNA was primed with a Not I - oligo(dT) primer [5'
                      TGTTACCAATCTGAAGTGGGAGCGGCCGCATCTTTTTTTTTTTTTTTTT 3'],
                      double-stranded cDNA was size selected, ligated to Eco RI
```

adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT7T3 vector
(Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by
M.Fatima Bonaldo. This library was constructed from the
same fetus as the fetal lung library, Soares fetal lung
NbHL19W."
/clone="356867"
/clone_lib="Soares fetal heart NbHH19W"
/sex="unknown"
/dev_stage="19 weeks"
/lab_host="DH10B (ampicillin resistant)"
mRNA complement(<1..>409)
BASE COUNT 130 a 96 c 98 g 85 t
ORIGIN

Query Match 1.9%; Score 24; DB 175; Length 409;
Best Local Similarity 69.4%; Pred. No. 4.59e-05;
Matches 43; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

```
Db    266 ggaaaacatcctgggtggtctgggctaaatgatcaagtgaaacagaaagccaggaaggat 325
          ||||||| ||||  |||  ||||||||  || ||    ||| ||||  ||  || ||| |
Cp    649 GGAAAACCTCCTCTGTGTCCTGGGCTACCTGCTCCTCAGAAGCAGAGGGCAGGGCAGGCT 590

Db    326 ct 327
          ||
Cp    589 CT 588
```

Search completed: Thu Aug 21 13:59:38 1997
Job time : 2124 secs.

****

MPSRCH (TM)

****

Release 2.1D John F. Collins, Biocomputing Research Unit.

MPsrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Thu Aug 21 14:00:02 1997; MasPar time 343.89 Seconds
1028.472 Million cell updates/sec

Tabular output not generated.

Title: >US-08-320-157-8
Description: (1-1287) from US08320157.seq
Perfect Score: 1287
N.A. Sequence: 1 TTTTAATATAAATTAATGTG.....CCTCAAGAGTACAGAAGCTT 1287
Comp: AAAATTATATTTAATTACAC.....GGAGTTCTCATGTCTTCGAA

Scoring table: TABLE default
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 359085 seqs, 137405154 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: EST-STS-THREE
1:EST199 2:EST200 3:EST201 4:EST202 5:EST203 6:EST204
7:EST205 8:EST206 9:EST207 10:EST208 11:EST209 12:EST210
13:EST211 14:EST212 15:EST213 16:EST214 17:EST215
18:EST216 19:EST217 20:EST218 21:EST219 22:EST220
23:EST221 24:EST222 25:EST223 26:EST224 27:EST225
28:EST226 29:EST227 30:EST228 31:EST229 32:EST230
33:EST231 34:EST232 35:EST233 36:STS1 37:STS2 38:STS3
39:STS4 40:STS5 41:STS6 42:STS7 43:STS8 44:STS9 45:STS10
46:STS11 47:STS12 48:STS13

Database: EST-STS-FOUR
49:gnEST1 50:gnEST2 51:gnEST3 52:gnEST4 53:gnEST5
54:gnEST6 55:gnEST7 56:gnEST8 57:gnEST9 58:gnEST10
59:gnEST11 60:gnEST12 61:gnEST13 62:gnEST14 63:gnEST15
64:gnEST16 65:gnEST17 66:gnEST18 67:gnEST19 68:gnEST20
69:gnEST21 70:gnEST22 71:gnEST23 72:gnEST24 73:gnEST25
74:gnSTS 75:enEST1 76:enEST2 77:enEST3 78:enEST4
79:enEST5 80:enEST6 81:enEST7 82:enEST8 83:enEST9
84:enEST10 85:enEST11 86:enEST12 87:enEST13 88:enEST14
89:enEST15 90:enEST16 91:enEST17 92:enEST18 93:enEST19
94:enEST20 95:enEST21 96:enEST22 97:enEST23 98:enEST24
99:enEST25 100:enEST26 101:enEST27 102:enEST28
103:enEST29 104:enEST30 105:enEST31 106:enSTS 107:ueEST1
108:ueEST2

Statistics: Mean 11.601; Variance 2.220; scale 5.227

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	% Query	No.	Score	Match	Length	DB	ID	Description	Pred. No.
		1	94	7.3	1258	38	G06794	human STS WI-7983.	5.87e-129
		2	55	4.3	133	91	MM1142620	mt62a08.r1 Soares 2Nb	2.88e-54
		3	55	4.3	133	61	AA185611	mt62a08.r1 Soares 2Nb	2.88e-54
		4	55	4.3	133	13	AA185611	mt62a08.r1 Soares 2Nb	2.88e-54
		5	40	3.1	310	7	AA123834	mp93c04.r1 Soares 2Nb	2.02e-28
		6	40	3.1	310	98	MMAA23834	mp93c04.r1 Soares 2Nb	2.02e-28
		7	40	3.1	310	58	AA123834	mp93c04.r1 Soares 2Nb	2.02e-28
		8	39	3.0	320	56	AA139013	mr04a06.r1 Soares mou	8.38e-27
		9	39	3.0	331	49	AA220617	my25f11.r1 Barstead m	8.38e-27
		10	39	3.0	331	108	MM1139075	my25f11.r1 Barstead m	8.38e-27
c		11	26	2.0	475	104	MMAA84020	mt22d08.r1 Soares mou	1.07e-07
c		12	26	2.0	475	59	AA184020	mt22d08.r1 Soares mou	1.07e-07
c		13	26	2.0	475	13	AA184020	mt22d08.r1 Soares mou	1.07e-07
c		14	24	1.9	228	1	N98063	2149C3 czapPFDd2.1, D	3.33e-05
		15	25	1.9	289	1	N98077	2232C3 czapPFDd2.1, D	1.98e-06
c		16	24	1.9	318	17	AA198152	mv49a10.r1 Soares mou	3.33e-05
c		17	24	1.9	412	25	N59408	yv51e02.s1 Soares fet	3.33e-05
c		18	24	1.9	442	8	AA126907	zl87b04.s1 Stratagene	3.33e-05
		19	24	1.9	457	15	AA191208	zq45f01.r1 Stratagene	3.33e-05
		20	24	1.9	553	4	AA115608	zk89a06.r1 Soares pre	3.33e-05
c		21	24	1.9	560	4	AA115474	zk89a06.s1 Soares pre	3.33e-05
c		22	23	1.8	99	25	N58693	yv64f10.r1 Soares fet	5.10e-04
c		23	23	1.8	250	40	G14786	human STS SHGC-13548.	5.10e-04
c		24	23	1.8	283	1	N97744	1186C3 czapPFDd2.1, D	5.10e-04
		25	23	1.8	301	56	AA138853	mr03a06.r1 Soares mou	5.10e-04
		26	23	1.8	311	104	MMAA89753	mt97c07.r1 Soares mou	5.10e-04
		27	23	1.8	311	59	AA189753	mt97c07.r1 Soares mou	5.10e-04
		28	23	1.8	311	15	AA189753	mt97c07.r1 Soares mou	5.10e-04
		29	23	1.8	320	64	AA231246	mw38h09.r1 Soares mou	5.10e-04
		30	23	1.8	320	99	MMAA31246	mw38h09.r1 Soares mou	5.10e-04
		31	23	1.8	327	24	HUMSUPY306	Human brain cDNA (fet	5.10e-04
		32	23	1.8	344	9	AA130593	zo12h07.s1 Stratagene	5.10e-04
c		33	23	1.8	372	40	G14445	human STS SHGC-10121.	5.10e-04
		34	23	1.8	385	93	MM1171296	mz84f08.r1 Soares mou	5.10e-04
		35	23	1.8	385	72	AA254866	mz84f08.r1 Soares mou	5.10e-04
		36	23	1.8	395	1	N97957	1718C3 czapPFDd2.1, D	5.10e-04
c		37	23	1.8	465	42	G21837	human STS WI-11990.	5.10e-04
		38	23	1.8	474	62	AA224645	CpEST.117 pSKIIminusC	5.10e-04
		39	23	1.8	474	80	CPAA24645	CpEST.117 pSKIIminusC	5.10e-04
c		40	23	1.8	492	6	AA121520	zk88g10.s1 Soares pre	5.10e-04
		41	22	1.7	286	12	AA182287	mt76h12.r1 Soares mou	6.99e-03
		42	22	1.7	297	79	BM4163	RRAMCA1275SK Brugia m	6.99e-03
c		43	22	1.7	349	17	AA197101	zq10f12.r1 Stratagene	6.99e-03
		44	22	1.7	364	71	AA250729	zr77f05.r1 Soares NhH	6.99e-03
c		45	22	1.7	534	104	MMAA85558	mt62e08.r1 Soares 2Nb	6.99e-03

ALIGNMENTS

RESULT 1
LOCUS G06794 1258 bp DNA STS 19-OCT-1995
DEFINITION human STS WI-7983.
ACCESSION G06794
NID g860039
KEYWORDS STS sequence; primer; sequence tagged site.
SOURCE human STSs derived from sequences in dbEST and the Unigene collection.
ORGANISM Homo sapiens
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Archonta; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1258)
AUTHORS Hudson,T.
TITLE Whitehead Institute/MIT Center for Genome Research; Physically Mapped ESTs
JOURNAL Unpublished (1995)
COMMENT
Contact: Thomas Hudson
Whitehead Institute/MIT Center for Genome Research
Whitehead Institute for Biomedical Research

```
          9 Cambridge Center, Cambridge MA 02142 USA
          Tel: 617 252 1900
          Fax: 617 252 1902
          Email: thudson@genome.wi.mit.edu

          Primer A: CTGATAACTTGGGGAGGCAA
          Primer B: GAGAGTCCAACTGCAAAGGC
          STS size: 345
          PCR Profile:
               Presoak:
               Denaturation:
               Annealing: 56 degrees C
               Polymerization:
               PCR Cycles: 35
               Thermal Cycler:
          Protocol:
              Template: 10 ng
              Primer: each 5 pM
              dNTPs: each 4 nM
              Taq Polymerase: 0.025 units/ul
              Total Vol: 20 ul

          Buffer:
              MgCl2: 1.5 mM
              KCl: 50 mM
              Tris-HCL: 10 mM
              pH: 9.3

          Prepared with primer pairs derived from U16811 -- Unigene.
FEATURES             Location/Qualifiers
     source          1..1258
                     /organism="Homo sapiens"
     STS             281..625
                     /map="751_A_7; 787_E_4; 908_C_6; 923_D_2; 919_D_10;
                     942_A_12; 808_G_(6,11)"
     primer_bind     281..300
                     /map="751_A_7; 787_E_4; 908_C_6; 923_D_2; 919_D_10;
                     942_A_12; 808_G_(6,11)"
     primer_bind     complement(606..625)
                     /map="751_A_7; 787_E_4; 908_C_6; 923_D_2; 919_D_10;
                     942_A_12; 808_G_(6,11)"
BASE COUNT      236 a    356 c    355 g    301 t     10 others
ORIGIN

  Query Match             7.3%;  Score 94;  DB 38;  Length 1258;
  Best Local Similarity  95.4%;  Pred. No. 5.87e-129;
  Matches   104;  Conservative    0;  Mismatches   4;  Indels   1;  Gaps   1;

Db      1 ctcccaagggtgccctttgggtcccggttcagacccctgcctggacttaagcgaagtctt 60
          ||||||||||||||| || ||||||| ||||||||||||||||||||||||||||||||||
Qy   1180 CTCCCAAGGGTGCCTTTGGGGTCCCAGTTCAGACCCCTGCCTGGACTTAAGCGAAGTCTT 1239

Db     61 tgccttctctgttcccttgcagggtccccccctcaagagtacagaagctt 109
          |||||||||||| ||| ||||||||||||||||||||||||||||||||
Qy   1240 TGCCTTCTCTGCTCC-TTGCAGGGTCCCCCCTCAAGAGTACAGAAGCTT 1287


RESULT    2
ID   MM1142620  standard; RNA; EST; 133 BP.
AC   AA185611;
NI   g1769266
DT   21-FEB-1997 (Rel. 51, Created)
DT   22-FEB-1997 (Rel. 51, Last updated, Version 2)
DE   mt62a08.r1 Soares 2NbMT Mus musculus cDNA clone 634454 5' similar
DE   to TR:G595926 G595926 BAK-2 PROTEIN. ;.
KW   EST.
OS   Mus musculus (house mouse)
OC   Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
OC   Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
OC   Mus.
RN   [1]
RP   1-133
RA   Marra M., Hillier L., Allen M., Bowles M., Dietrich N.,
RA   Dubuque T., Geisel S., Kucaba T., Lacy M., Le M., Martin J.,
RA   Morris M., Schellenberg K., Steptoe M., Tan F., Underwood K.,
RA   Moore B., Theising B., Wylie T., Lennon G., Soares B., Wilson R.,
RA   Waterston R.;
RT   "The WashU-HHMI Mouse EST Project";
RL   Unpublished.
CC   Contact: Marra M/Mouse EST Project WashU-HHMI Mouse EST Project
CC   Washington University School of MedicineP 4444 Forest Park Parkway,
CC   Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810
CC   Email: mouseest@watson.wustl.edu This clone is available
CC   royalty-free through LLNL ; contact the IMAGE Consortium
CC   (info@image.llnl.gov) for further information. MGI:386446 Trace
CC   considered overall poor quality Possible reversed clone: similarity
CC   on wrong strand Seq primer: -28M13 rev2 from Amersham High quality
CC   sequence stop: 1.
FH   Key             Location/Qualifiers
FH
FT   source          1..133
FT                   /organism="Mus musculus"
FT                   /strain="C57BL/6J"
FT                   /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
FT                   polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
FT                   was primed with a Not I - oligo(dT) primer [5'
FT                   TGTTACCAATCTGAAGTGGGAGCGGCCGCGTTTTTTTTTTTTTTTTTTTTTTT
FT                   3']; double-stranded cDNA wasligated to Eco RI adaptors
FT                   (Pharmacia), digested with Not I and cloned into the Not I
FT                   and Eco RI sites of the modified pT7T3 vector. RNA provide
  d
FT                   by Dr. Bertrand Jordan. Library went through two rounds of
FT                   normalization, and was constructedby Bento Soares and
FT                   M.Fatima Bonaldo."
FT                   /clone="634454"
FT                   /clone_lib="Soares 2NbMT"
FT                   /sex="male"
FT                   /tissue_type="Thymus"
FT                   /dev_stage="4 weeks"
FT                   /lab_host="DH10B"
FT   mRNA            <1..>133
SQ   Sequence 133 BP; 33 A; 40 C; 34 G; 26 T; 0 other;

  Query Match             4.3%;  Score 55;  DB 91;  Length 133;
  Best Local Similarity  81.6%;  Pred. No. 2.88e-54;
  Matches    71;  Conservative    0;  Mismatches  16;  Indels   0;  Gaps   0;

Db     41 ccgtccccttctggacagcaggttgcccatgacacagaggaggtctttcgaagctacgta 100
          || ||  ||||||   ||||||| ||||| ||||||||||||||| || || ||||||||
Qy    601 CCCTCTGCTTCTGAGGAGCAGGTAGCCCAGGACACAGAGGAGGTTTTCCGCAGCTACGTT 660

Db    101 ttttaccttcaccagcaggaacatgag 127
          |||||||  || ||||||||||| |||
Qy    661 TTTTACCACCATCAGCAGGAACAGGAG 687


RESULT    3
LOCUS       AA185611      133 bp    mRNA            EST       19-FEB-1997
DEFINITION  mt62a08.r1 Soares 2NbMT Mus musculus cDNA clone 634454 5' similar
            to TR:G595926 G595926 BAK-2 PROTEIN. ;.
ACCESSION   AA185611
NID         g1769266
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Mus.
REFERENCE   1  (bases 1 to 133)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
```

```
COMMENT
          Contact: Marra M/Mouse EST Project
          WashU-HHMI Mouse EST Project
          Washington University School of MedicineP
          4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
          Tel: 314 286 1800
          Fax: 314 286 1810
          Email: mouseest@watson.wustl.edu
          This clone is available royalty-free through LLNL ; contact the
          IMAGE Consortium (info@image.llnl.gov) for further information.
          MGI:386446
          Trace considered overall poor quality
          Possible reversed clone: similarity on wrong strand
          Seq primer: -28M13 rev2 from Amersham
          High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1..133
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGTTTTTTTTTTTTTTTTTTTTTTTT
                     3']; double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. RNA
                     provided by Dr. Bertrand Jordan. Library went through two
                     rounds of normalization, and was constructed by Bento
                     Soares and M.Fatima Bonaldo."
                     /clone="634454"
                     /clone_lib="Soares 2NbMT"
                     /sex="male"
                     /tissue_type="Thymus"
                     /dev_stage="4 weeks"
                     /lab_host="DH10B"
     mRNA            <1..>133
BASE COUNT       33 a     40 c     34 g     26 t
ORIGIN

  Query Match             4.3%;  Score 55;  DB 61;  Length 133;
  Best Local Similarity  81.6%;  Pred. No. 2.88e-54;
  Matches    71;  Conservative    0;  Mismatches  16;  Indels   0;  Gaps   0;

Db     41 ccgtccccttctggacagcaggttgcccatgacacagaggaggtctttcgaagctacgta 100
          || ||  ||||||   ||||||| ||||| ||||||||||||||| || || ||||||||
Qy    601 CCCTCTGCTTCTGAGGAGCAGGTAGCCCAGGACACAGAGGAGGTTTTCCGCAGCTACGTT 660

Db    101 ttttaccttcaccagcaggaacatgag 127
          |||||||  || ||||||||||| |||
Qy    661 TTTTACCACCATCAGCAGGAACAGGAG 687


RESULT    4
LOCUS       AA185611      133 bp    mRNA            EST       07-JAN-1997
DEFINITION  mt62a08.r1 Soares 2NbMT Mus musculus cDNA clone 634454 5' similar
            to TR:G595926 G595926 BAK-2 PROTEIN. ;.
ACCESSION   AA185611
NID         g1769266
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Mus.
REFERENCE   1  (bases 1 to 133)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT
          Contact: Marra M/Mouse EST Project
          WashU-HHMI Mouse EST Project
          Washington University School of MedicineP
          4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
          Tel: 314 286 1800
          Fax: 314 286 1810
          Email: mouseest@watson.wustl.edu
          This clone is available royalty-free through LLNL ; contact the
          IMAGE Consortium (info@image.llnl.gov) for further information.
          MGI:386446
          Trace considered overall poor quality
          Possible reversed clone: similarity on wrong strand
          Seq primer: -28M13 rev2 from Amersham
          High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1..133
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGTTTTTTTTTTTTTTTTTTTTTTTT
                     3']; double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not
                     I and Eco RI sites of the modified pT7T3 vector. RNA
                     provided by Dr. Bertrand Jordan. Library went through two
                     rounds of normalization, and was constructed by Bento
                     Soares and M.Fatima Bonaldo."
                     /clone="634454"
                     /clone_lib="Soares 2NbMT"
                     /sex="male"
                     /dev_stage="4 weeks"
                     /lab_host="DH10B"
     mRNA            <1..>133
BASE COUNT       33 a     40 c     34 g     26 t
ORIGIN

  Query Match             4.3%;  Score 55;  DB 13;  Length 133;
  Best Local Similarity  81.6%;  Pred. No. 2.88e-54;
  Matches    71;  Conservative    0;  Mismatches  16;  Indels   0;  Gaps   0;

Db     41 ccgtccccttctggacagcaggttgcccatgacacagaggaggtctttcgaagctacgta 100
          || ||  ||||||   ||||||| ||||| ||||||||||||||| || || ||||||||
Qy    601 CCCTCTGCTTCTGAGGAGCAGGTAGCCCAGGACACAGAGGAGGTTTTCCGCAGCTACGTT 660

Db    101 ttttaccttcaccagcaggaacatgag 127
          |||||||  || ||||||||||| |||
Qy    661 TTTTACCACCATCAGCAGGAACAGGAG 687


RESULT    5
LOCUS       AA123834      310 bp    mRNA            EST       21-NOV-1996
DEFINITION  mp93c04.r1 Soares 2NbMT Mus musculus cDNA clone 576774 5' similar
            to TR:G595926 G595926 BAK-2 PROTEIN. ;.
ACCESSION   AA123834
NID         g1682242
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Mus.
REFERENCE   1  (bases 1 to 310)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT
          Contact: Marra M/Mouse EST Project
          WashU-HHMI Mouse EST Project
          Washington University School of MedicineP
```

```
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:351422
            Seq primer: -28M13 rev2 from Amersham.
FEATURES              Location/Qualifiers
     source           1..310
                      /organism="Mus musculus"
                      /strain="C57BL/6J"
                      /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                      polylinker; Site_1: Not I; Site_2: Eco RI; [5'
                      TGTTACCAATCTGAAGTGGGAGCGGCCGCGTTTTTTTTTTTTTTTTTTTTTTT
                      3']; double-stranded cDNA was ligated to Eco RI adaptors
                      (Pharmacia), digested with Not I and cloned into the Not
                      I and Eco RI sites of the modified pT7T3 vector. RNA
                      provided by Dr. Bertrand Jordan. Library went through two
                      rounds of normalization, and was constructed by Bento
                      Soares and M.Fatima Bonaldo."
                      /clone="576774"
                      /clone_lib="Soares 2NbMT"
                      /sex="male"
                      /dev_stage="4 weeks"
                      /lab_host="DH10B"
     mRNA             <1..>310
BASE COUNT       72 a     96 c     86 g     55 t      1 others
ORIGIN

  Query Match             3.1%;  Score 40;  DB 7;  Length 310;
  Best Local Similarity  78.2%;  Pred. No. 2.02e-28;
  Matches   104;  Conservative    0;  Mismatches  22;  Indels   7;  Gaps   5;

Db    185 ctccaccaagacctgaaaaatggcatctggacaaggaccaggtcccccgaagtgggc-tg 243
          ||||||  |||||||||||||||| || || ||||| |||||||| || | |  ||  ||
Qy    525 CTCCACTGAGACCTGAAAAATGGCTTCGGGGCAAGGCCCAGGTCCTCCCAGGCAGGAGTG 584

Db    244 cgatgagtc--ccc-gtcc-c--cttctgaacagcaggttgcccaggacacagaggaggt 297
          ||  ||| |  ||| | || |  |||||||  ||||||| ||||||||||||||||||||
Qy    585 CGGAGAGCCTGCCCTGCCCTCTGCTTCTGAGGAGCAGGTAGCCCAGGACACAGAGGAGGT 644

Db    298 ctttcgaagctac 310
          || || ||||||
Qy    645 TTTCCGCAGCTAC 657


RESULT    6
ID    MMAA23834  standard; RNA; EST; 310 BP.
AC    AA123834;
NI    g1682242
DT    25-NOV-1996 (Rel. 50, Created)
DT    19-FEB-1997 (Rel. 51, Last updated, Version 2)
DE    mp93c04.r1 Soares 2NbMT Mus musculus cDNA clone 576774 5' similar
DE    to TR:G595926 G595926 BAK-2 PROTEIN. ;.
KW    EST.
OS    Mus musculus (house mouse)
OC    Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
OC    Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
OC    Mus.
RN    [1]
RP    1-310
RA    Marra M., Hillier L., Allen M., Bowles M., Dietrich N.,
RA    Dubuque T., Geisel S., Kucaba T., Lacy M., Le M., Martin J.,
RA    Morris M., Schellenberg K., Steptoe M., Tan F., Underwood K.,
RA    Moore B., Theising B., Wylie T., Lennon G., Soares B., Wilson R.,
RA    Waterston R.;
RT    "The WashU-HHMI Mouse EST Project";
RL    Unpublished.
CC    Contact: Marra M/Mouse EST Project WashU-HHMI Mouse EST Project
CC    Washington University School of MedicineP 4444 Forest Park Parkway,
CC    Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810
CC    Email: mouseest@watson.wustl.edu This clone is available
CC    royalty-free through LLNL ; contact the IMAGE Consortium
CC    (info@image.llnl.gov) for further information. MGI:351422 Seq
CC    primer: -28M13 rev2 from Amersham.
FH    Key             Location/Qualifiers
FH
FT    source          1..310
FT                    /organism="Mus musculus"
FT                    /strain="C57BL/6J"
FT                    /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
FT                    polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
FT                    was primed with a Not I - oligo(dT) primer [5'
FT                    TGTTACCAATCTGAAGTGGGAGCGGCCGCGTTTTTTTTTTTTTTTTTTTTTTT
FT                    3']; double-stranded cDNA wasligated to Eco RI adaptors
FT                    (Pharmacia), digested with Not I and cloned into the Not I
FT                    and Eco RI sites of the modified pT7T3 vector. RNA provide
  d
FT                    by Dr. Bertrand Jordan. Library went through two rounds of
FT                    normalization, and was constructedby Bento Soares and
FT                    M.Fatima Bonaldo."
FT                    /clone="576774"
FT                    /clone_lib="Soares 2NbMT"
FT                    /sex="male"
FT                    /tissue_type="Thymus"
FT                    /dev_stage="4 weeks"
FT                    /lab_host="DH10B"
FT    mRNA            <1..>310
SQ    Sequence 310 BP; 72 A; 96 C; 86 G; 55 T; 1 other;

  Query Match             3.1%;  Score 40;  DB 98;  Length 310;
  Best Local Similarity  78.2%;  Pred. No. 2.02e-28;
  Matches   104;  Conservative    0;  Mismatches  22;  Indels   7;  Gaps   5;

Db    185 ctccaccaagacctgaaaaatggcatctggacaaggaccaggtcccccgaagtgggc-tg 243
          ||||||  |||||||||||||||| || || ||||| |||||||| || | |  ||  ||
Qy    525 CTCCACTGAGACCTGAAAAATGGCTTCGGGGCAAGGCCCAGGTCCTCCCAGGCAGGAGTG 584

Db    244 cgatgagtc--ccc-gtcc-c--cttctgaacagcaggttgcccaggacacagaggaggt 297
          ||  ||| |  ||| | || |  |||||||  ||||||| ||||||||||||||||||||
Qy    585 CGGAGAGCCTGCCCTGCCCTCTGCTTCTGAGGAGCAGGTAGCCCAGGACACAGAGGAGGT 644

Db    298 ctttcgaagctac 310
          || || ||||||
Qy    645 TTTCCGCAGCTAC 657


RESULT    7
LOCUS       AA123834      310 bp    mRNA            EST       17-FEB-1997
DEFINITION  mp93c04.r1 Soares 2NbMT Mus musculus cDNA clone 576774 5' similar
            to TR:G595926 G595926 BAK-2 PROTEIN. ;.
ACCESSION   AA123834
NID         g1682242
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Mus.
REFERENCE   1  (bases 1 to 310)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT
            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
```

```
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:351422
            Seq primer: -28M13 rev2 from Amersham.
FEATURES              Location/Qualifiers
     source           1..310
                      /organism="Mus musculus"
                      /strain="C57BL/6J"
                      /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                      polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                      was primed with a Not I - oligo(dT) primer [5'
                      TGTTACCAATCTGAAGTGGGAGCGGCCGCGTTTTTTTTTTTTTTTTTTTTTT
                      3']; double-stranded cDNA was ligated to Eco RI adaptors
                      (Pharmacia), digested with Not I and cloned into the Not I
                      and Eco RI sites of the modified pT7T3 vector. RNA
                      provided by Dr. Bertrand Jordan. Library went through two
                      rounds of normalization, and was constructed by Bento
                      Soares and M.Fatima Bonaldo."
                      /clone="576774"
                      /clone_lib="Soares 2NbMT"
                      /sex="male"
                      /tissue_type="Thymus"
                      /dev_stage="4 weeks"
                      /lab_host="DH10B"
     mRNA             <1..>310
BASE COUNT       72 a     96 c     86 g     55 t      1 others
ORIGIN

  Query Match             3.1%;  Score 40;  DB 58;  Length 310;
  Best Local Similarity  78.2%;  Pred. No. 2.02e-28;
  Matches   104;  Conservative    0;  Mismatches  22;  Indels   7;  Gaps   5;

Db    185 ctccaccaagacctgaaaaatggcatctggacaaggaccaggtcccccgaagtgggc-tg 243
          ||||||  ||||||||||||||| || || ||||| |||||||| || | |  ||  ||
Qy    525 CTCCACTGAGACCTGAAAAATGGCTTCGGGGCAAGGCCCAGGTCCTCCCAGGCAGGAGTG 584

Db    244 cgatgagtc--ccc-gtcc-c--cttctgaacagcaggttgcccaggacacagaggaggt 297
          ||  ||| |  ||| | || |  |||||||  ||||||| ||||||||||||||||||||
Qy    585 CGGAGAGCCTGCCCTGCCCTCTGCTTCTGAGGAGCAGGTAGCCCAGGACACAGAGGAGGT 644

Db    298 ctttcgaagctac 310
            || || ||||||
Qy    645 TTTCCGCAGCTAC 657


RESULT    8
LOCUS       AA139013      320 bp    mRNA            EST       16-FEB-1997
DEFINITION  mr04a06.r1 Soares mouse 3NbMS Mus musculus cDNA clone 596434 5'
            similar to TR:G595926 G595926 BAK-2 PROTEIN. ;.
ACCESSION   AA139013
NID         g1701240
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Mus.
REFERENCE   1  (bases 1 to 320)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT
            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:361866
            Seq primer: -28M13 rev2 from Amersham.
FEATURES              Location/Qualifiers
     source           1..320
                      /organism="Mus musculus"
                      /strain="C57BL/6J"
                      /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                      polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                      was primed with a Not I - oligo(dT) primer [5'
                      TGTTACCAATCTGAAGTGGGAGCGGCCGCGCTGTTTTTTTTTTTTTTTTTTTTT
                      3']; double-stranded cDNA was ligated to Eco RI adaptors
                      (Pharmacia), digested with Not I and cloned into the Not I
                      and Eco RI sites of the modified pT7T3 vector. RNA
                      provided by Dr. Bertrand Jordan. Library went through
                      three rounds of normalization, and was constructed by
                      Bento Soares and M.Fatima Bonaldo."
                      /clone="596434"
                      /clone_lib="Soares mouse 3NbMS"
                      /sex="male"
                      /tissue_type="Spleen"
                      /dev_stage="4 weeks"
                      /lab_host="DH10B"
     mRNA             <1..>320
BASE COUNT       68 a     92 c     96 g     64 t
ORIGIN

  Query Match             3.0%;  Score 39;  DB 56;  Length 320;
  Best Local Similarity  84.2%;  Pred. No. 8.38e-27;
  Matches    48;  Conservative    0;  Mismatches   9;  Indels   0;  Gaps   0;

Db    264 ccgtccccttctgaacagcaggttgcccaggacacagaggaggtctttcgaagctac 320
          || ||  |||||||  ||||||| |||||||||||||||||||| || || ||||||
Qy    601 CCCTCTGCTTCTGAGGAGCAGGTAGCCCAGGACACAGAGGAGGTTTTCCGCAGCTAC 657


RESULT    9
LOCUS       AA220617      331 bp    mRNA            EST       10-FEB-1997
DEFINITION  my25f11.r1 Barstead mouse pooled organs MPLRB4 Mus musculus cDNA
            clone 696909 5'.
ACCESSION   AA220617
NID         g1838404
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Mus.
REFERENCE   1  (bases 1 to 331)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT
            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:430469
            Seq primer: -28m13 rev2 ET from Amersham.
FEATURES              Location/Qualifiers
     source           1..331
                      /organism="Mus musculus"
                      /strain="FVB/N"
```

```
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: EcoRI; Site_2: NotI; 1st strand cDNA
                     was primed with a Not I - oligo(dT) primer [5'
                     TGTTACGAATCTGAAGTGGGAGCGGCCGCCCTTTTTTTTTTTTTTTTTTTTTTT
                     3']; double-stranded cDNA was ligated to Eco RI adaptors
                     [GTTGGATTCGGTACC], digested with Not I and cloned into
                     the Not I and Eco RI sites of the modified pT7T3 vector.
                     Library constructed by Bob Barstead."
                     /clone="696909"
                     /clone_lib="Barstead mouse pooled organs MPLRB4"
                     /sex="mixed"
                     /dev_stage="7 day"
                     /lab_host="DH10B"
     mRNA            <1..>331
BASE COUNT       78 a    104 c     89 g     60 t
ORIGIN

  Query Match            3.0%;  Score 39;  DB 49;  Length 331;
  Best Local Similarity  84.2%;  Pred. No. 8.38e-27;
  Matches    48;  Conservative    0;  Mismatches   9;  Indels   0;  Gaps   0;

Db    275 ccgtccccttctgaacagcaggttgcccaggacacagaggaggtctttcgaagctac 331
          || ||  |||||||  ||||||| |||||||||||||||||||| || || ||||||
Qy    601 CCCTCTGCTTCTGAGGAGCAGGTAGCCCAGGACACAGAGGAGGTTTTCCGCAGCTAC 657


RESULT   10
ID   MM1139075  standard; RNA; EST; 331 BP.
AC   AA220617;
NI   g1838404
DT   13-FEB-1997 (Rel. 50, Created)
DT   13-FEB-1997 (Rel. 50, Last updated, Version 1)
DE   my25f11.r1 Barstead mouse pooled organs MPLRB4 Mus musculus cDNA
DE   clone 696909 5'.
KW   EST.
OS   Mus musculus (house mouse)
OC   Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
OC   Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
OC   Mus.
RN   [1]
RP   1-331
RA   Marra M., Hillier L., Allen M., Bowles M., Dietrich N.,
RA   Dubuque T., Geisel S., Kucaba T., Lacy M., Le M., Martin J.,
RA   Morris M., Schellenberg K., Steptoe M., Tan F., Underwood K.,
RA   Moore B., Theising B., Wylie T., Lennon G., Soares B., Wilson R.,
RA   Waterston R.;
RT   "The WashU-HHMI Mouse EST Project";
RL   Unpublished.
CC   Contact: Marra M/Mouse EST Project WashU-HHMI Mouse EST Project
CC   Washington University School of MedicineP 4444 Forest Park Parkway,
CC   Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810
CC   Email: mouseest@watson.wustl.edu This clone is available
CC   royalty-free through LLNL ; contact the IMAGE Consortium
CC   (info@image.llnl.gov) for further information. MGI:430469 Seq
CC   primer: -28m13 rev2 ET from Amersham.
FH   Key             Location/Qualifiers
FH
FT   source          1..331
FT                   /organism="Mus musculus"
FT                   /strain="FVB/N"
FT                   /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
FT                   polylinker; Site_1: EcoRI; Site_2: NotI; 1st strand cDNA
FT                   was primed with a Not I - oligo(dT) primer [5'
FT                   TGTTACGAATCTGAAGTGGGAGCGGCCGCCCTTTTTTTTTTTTTTTTTTTTTTT
FT                   3']; double-stranded cDNA wasligated to Eco RI adaptors
FT                   [GTTGGATTCGGTACC], digested with Not I and cloned into the
FT                   Not I and Eco RI sites of the modified pT7T3 vector.
FT                   Library constructed by Bob Barstead."
FT                   /clone="696909"
FT                   /clone_lib="Barstead mouse pooled organs MPLRB4"
FT                   /sex="mixed"
FT                   /dev_stage="7 day"
FT                   /lab_host="DH10B"
FT   mRNA            <1..>331
SQ   Sequence 331 BP; 78 A; 104 C; 89 G; 60 T; 0 other;

  Query Match            3.0%;  Score 39;  DB 108;  Length 331;
  Best Local Similarity  84.2%;  Pred. No. 8.38e-27;
  Matches    48;  Conservative    0;  Mismatches   9;  Indels   0;  Gaps   0;

Db    275 ccgtccccttctgaacagcaggttgcccaggacacagaggaggtctttcgaagctac 331
          || ||  |||||||  ||||||| |||||||||||||||||||| || || ||||||
Qy    601 CCCTCTGCTTCTGAGGAGCAGGTAGCCCAGGACACAGAGGAGGTTTTCCGCAGCTAC 657


RESULT   11
ID   MMAA84020  standard; RNA; EST; 475 BP.
AC   AA184020;
NI   g1767585
DT   19-FEB-1997 (Rel. 51, Created)
DT   19-FEB-1997 (Rel. 51, Last updated, Version 1)
DE   mt22d08.r1 Soares mouse 3NbMS Mus musculus cDNA clone 621807 5'.
KW   EST.
OS   Mus musculus (house mouse)
OC   Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
OC   Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
OC   Mus.
RN   [1]
RP   1-475
RA   Marra M., Hillier L., Allen M., Bowles M., Dietrich N.,
RA   Dubuque T., Geisel S., Kucaba T., Lacy M., Le M., Martin J.,
RA   Morris M., Schellenberg K., Steptoe M., Tan F., Underwood K.,
RA   Moore B., Theising B., Wylie T., Lennon G., Soares B., Wilson R.,
RA   Waterston R.;
RT   "The WashU-HHMI Mouse EST Project";
RL   Unpublished.
CC   Contact: Marra M/Mouse EST Project WashU-HHMI Mouse EST Project
CC   Washington University School of MedicineP 4444 Forest Park Parkway,
CC   Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810
CC   Email: mouseest@watson.wustl.edu This clone is available
CC   royalty-free through LLNL ; contact the IMAGE Consortium
CC   (info@image.llnl.gov) for further information. MGI:382631 Seq
CC   primer: -28M13 rev2 from Amersham High quality sequence stop: 438.
FH   Key             Location/Qualifiers
FH
FT   source          1..475
FT                   /organism="Mus musculus"
FT                   /strain="C57BL/6J"
FT                   /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
FT                   polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
FT                   was primed with a Not I - oligo(dT) primer [5'
FT                   TGTTACCAATCTGAAGTGGGAGCGGCCGCGCTGTTTTTTTTTTTTTTTTTTTTTT
FT                   3']; double-stranded cDNA was ligated to Eco RI adaptors
FT                   (Pharmacia), digested with Not I and cloned into the Not I
FT                   and Eco RI sites of the modified pT7T3 vector. RNA provide
  d
FT                   by Dr. Bertrand Jordan. Library went through three rounds
FT                   of normalization, and was constructed by Bento Soares and
FT                   M.Fatima Bonaldo."
FT                   /clone="621807"
FT                   /clone_lib="Soares mouse 3NbMS"
FT                   /sex="male"
FT                   /tissue_type="Spleen"
FT                   /dev_stage="4 weeks"
FT                   /lab_host="DH10B"
FT   mRNA            <1..>475
SQ   Sequence 475 BP; 150 A; 56 C; 67 G; 202 T; 0 other;

  Query Match            2.0%;  Score 26;  DB 104;  Length 475;
  Best Local Similarity  76.0%;  Pred. No. 1.07e-07;
  Matches    38;  Conservative    0;  Mismatches  12;  Indels   0;  Gaps   0;

Db    291 ttctaatatttttcatttctgtaacatttaaacttcctttgagtaaattt 340
          || | || ||||||| |||| |||||||| || ||   || |||| ||||
Cp     94 TTTTCATCTTTTTCAGTTCTATAACATTTGAATTTTTATTAAGTATATTT 45
```

```
RESULT   12
LOCUS       AA184020      475 bp    mRNA            EST       17-FEB-1997
DEFINITION  mt22d08.r1 Soares mouse 3NbMS Mus musculus cDNA clone 621807 5'.
ACCESSION   AA184020
NID         g1767585
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Mus.
REFERENCE   1  (bases 1 to 475)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT
            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:382631
            Seq primer: -28M13 rev2 from Amersham
            High quality sequence stop: 438.
FEATURES             Location/Qualifiers
     source          1..475
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGCTGTTTTTTTTTTTTTTTTTTTTTTT
                     3']; double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. RNA
                     provided by Dr. Bertrand Jordan. Library went through
                     three rounds of normalization, and was constructed by
                     Bento Soares and M.Fatima Bonaldo."
                     /clone="621807"
                     /clone_lib="Soares mouse 3NbMS"
                     /sex="male"
                     /tissue_type="Spleen"
                     /dev_stage="4 weeks"
                     /lab_host="DH10B"
     mRNA            <1..>475
BASE COUNT      150 a     56 c     67 g    202 t
ORIGIN

  Query Match             2.0%;  Score 26;  DB 59;  Length 475;
  Best Local Similarity  76.0%;  Pred. No. 1.07e-07;
  Matches    38;  Conservative    0;  Mismatches  12;  Indels   0;  Gaps   0;

Db    291 ttctaatatttttcatttctgtaacatttaaacttcctttgagtaaattt 340
          || | || ||||||| |||| |||||||| || ||   || |||| ||||
Cp     94 TTTTCATCTTTTTCAGTTCTATAACATTTGAATTTTTATTAAGTATATTT 45


RESULT   13
LOCUS       AA184020      475 bp    mRNA            EST       07-JAN-1997
DEFINITION  mt22d08.r1 Soares mouse 3NbMS Mus musculus cDNA clone 621807 5'.
ACCESSION   AA184020
NID         g1767585
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Mus.
REFERENCE   1  (bases 1 to 475)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT
            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:382631
            Seq primer: -28M13 rev2 from Amersham
            High quality sequence stop: 438.
FEATURES             Location/Qualifiers
     source          1..475
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGCTGTTTTTTTTTTTTTTTTTTTTTTT
                     3']; double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not
                     I and Eco RI sites of the modified pT7T3 vector. RNA
                     provided by Dr. Bertrand Jordan. Library went through
                     three rounds of normalization, and was constructed by
                     Bento Soares and M.Fatima Bonaldo."
                     /clone="621807"
                     /clone_lib="Soares mouse 3NbMS"
                     /sex="male"
                     /dev_stage="4 weeks"
                     /lab_host="DH10B"
     mRNA            <1..>475
BASE COUNT      150 a     56 c     67 g    202 t
ORIGIN

  Query Match             2.0%;  Score 26;  DB 13;  Length 475;
  Best Local Similarity  76.0%;  Pred. No. 1.07e-07;
  Matches    38;  Conservative    0;  Mismatches  12;  Indels   0;  Gaps   0;

Db    291 ttctaatatttttcatttctgtaacatttaaacttcctttgagtaaattt 340
          || | || ||||||| |||| |||||||| || ||   || |||| ||||
Cp     94 TTTTCATCTTTTTCAGTTCTATAACATTTGAATTTTTATTAAGTATATTT 45


RESULT   14
LOCUS       N98063        228 bp    mRNA            EST       18-NOV-1996
DEFINITION  2149C3 czapPFDd2.1, Debopam Chakrabarti Plasmodium falciparum cDNA
            clone PF2149C.
ACCESSION   N98063
NID         g1675113
KEYWORDS    EST.
SOURCE      malaria parasite.
  ORGANISM  Plasmodium falciparum
            Eukaryotae; mitochondrial eukaryotes; Alveolata; Apicomplexa;
            Haemosporida; Plasmodium.
REFERENCE   1  (bases 1 to 228)
  AUTHORS   Dame,J.B., Arnot,D.E., Bourke,P., Chakrabarti,D., Christodoulou,Z.,
            Coppel,R., Cowman,A., Craig,A., Fischer,K., Foster,J., Goodman,N.,
            Hinterberg,K., Holder,A.A., Holt,D., Kemp,D., Lanzer,M., Lim,A.,
            Newbold,C., Ravetch,J.V., Reddy,G.R., Rubio,J., Schuster,S.M.,
            Su,X.-Z., Thompson,J.K., Vital,F., Wellems,T.E. and Werner,E.
```

```
  TITLE     Current status of the Plasmodium falciparum genome project
  JOURNAL   Mol. Biochem. Parasitol. 79, 1-12 (1996)
  MEDLINE   97001675
COMMENT
            Contact: Debopam Chakrabarti
            Department of Molecular Biology and Microbiology
            University of Central Florida
            Orlando, FL 32816-2360
            Tel: 407 384 2061
            Fax: 407 384 3095
            Email: dchak@pegasus.cc.ucf.edu
            Seq primer: T3.
FEATURES             Location/Qualifiers
     source          1..228
                     /organism="Plasmodium falciparum"
                     /strain="Dd2"
                     /note="Vector: Lambda ZAP II; Site_1: EcoR I; Site_2: Xho
                     I; PolyA+ RNA, from asynchronous blood stage parasites of
                     the Dd2 isolate cultured in vitro, was reverse transcribed
                     using an oligo dT-Xho I primer.  Second strand was
                     prepared using RNase H and DNA polymerase I. EcoR I
                     adapters were ligated to the cDNA, and it was digested
                     with Xho I. Prepared fragments were ligated into EcoR I +
                     Xho I digested lambda ZAP II vector."
                     /clone="PF2149C"
                     /clone_lib="czapPFDd2.1, Debopam Chakrabarti"
                     /lab_host="E. coli XL-1 blue"
     mRNA            <1..>228
BASE COUNT       82 a     14 c      9 g    123 t
ORIGIN

  Query Match             1.9%;  Score 24;  DB 1;  Length 228;
  Best Local Similarity  68.8%;  Pred. No. 3.33e-05;
  Matches    44;  Conservative    0;  Mismatches  20;  Indels   0;  Gaps   0;

Db     48 gttttttttatattttaaattttaaagttttatttctttttatatttttatttttatttaatt 107
          ||| ||||||  ||| |  ||||  |||| |||  | |||  |||||  || | |  |||
Cp    105 GTTGTTTTTACTTTTCATCTTTTTCAGTTCTATAACATTTGAATTTTTATTAAGTATATT 46

Db    108 tcat 111
          ||||
Cp     45 TCAT 42


RESULT   15
LOCUS       N98077        289 bp    mRNA            EST       18-NOV-1996
DEFINITION  2232C3 czapPFDd2.1, Debopam Chakrabarti Plasmodium falciparum cDNA
            clone PF2232C.
ACCESSION   N98077
NID         g1675127
KEYWORDS    EST.
SOURCE      malaria parasite.
  ORGANISM  Plasmodium falciparum
            Eukaryotae; mitochondrial eukaryotes; Alveolata; Apicomplexa;
            Haemosporida; Plasmodium.
REFERENCE   1  (bases 1 to 289)
  AUTHORS   Dame,J.B., Arnot,D.E., Bourke,P., Chakrabarti,D., Christodoulou,Z.,
            Coppel,R., Cowman,A., Craig,A., Fischer,K., Foster,J., Goodman,N.,
            Hinterberg,K., Holder,A.A., Holt,D., Kemp,D., Lanzer,M., Lim,A.,
            Newbold,C., Ravetch,J.V., Reddy,G.R., Rubio,J., Schuster,S.M.,
            Su,X.-Z., Thompson,J.K., Vital,F., Wellems,T.E. and Werner,E.
  TITLE     Current status of the Plasmodium falciparum genome project
  JOURNAL   Mol. Biochem. Parasitol. 79, 1-12 (1996)
  MEDLINE   97001675
COMMENT
            Contact: Debopam Chakrabarti
            Department of Molecular Biology and Microbiology
            University of Central Florida
            Orlando, FL 32816-2360
            Tel: 407 384 2061
            Fax: 407 384 3095
            Email: dchak@pegasus.cc.ucf.edu
            Seq primer: T3.
FEATURES             Location/Qualifiers
     source          1..289
                     /organism="Plasmodium falciparum"
                     /strain="Dd2"
                     /note="Vector: Lambda ZAP II; Site_1: EcoR I; Site_2: Xho
                     I; PolyA+ RNA, from asynchronous blood stage parasites of
                     the Dd2 isolate cultured in vitro, was reverse transcribed
                     using an oligo dT-Xho I primer.  Second strand was
                     prepared using RNase H and DNA polymerase I. EcoR I
                     adapters were ligated to the cDNA, and it was digested
                     with Xho I. Prepared fragments were ligated into EcoR I +
                     Xho I digested lambda ZAP II vector."
                     /clone="PF2232C"
                     /clone_lib="czapPFDd2.1, Debopam Chakrabarti"
                     /lab_host="E. coli XL-1 blue"
     mRNA            <1..>289
BASE COUNT      138 a     16 c     15 g    119 t      1 others
ORIGIN

  Query Match             1.9%;  Score 25;  DB 1;  Length 289;
  Best Local Similarity  66.7%;  Pred. No. 1.98e-06;
  Matches    50;  Conservative    0;  Mismatches  25;  Indels   0;  Gaps   0;

Db    208 taaacacaaaaaattatatatatatattataaatataacatttgtataattaaaaaaatt 267
          || | |||| | || | |||||  || || ||||  ||   || || || | |||||  |
Qy     30 TAGAGACAATACATGAAATATACTTAATAAAAATTCAAATGTTATAGAACTGAAAAAGAT 89

Db    268 aaaaaaaaaaaaaaa 282
           ||||  ||||| ||
Qy     90 GAAAAGTAAAAACAA 104


Search completed: Thu Aug 21 14:17:25 1997
Job time : 1043 secs.
```

****

MPSRCH (TM)

****

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Wed Aug 20 11:04:54 1997; MasPar time 5.42 Seconds
445.465 Million cell updates/sec

Tabular output not generated.

Title: >US-08-320-157-9
Description: (1-211) from US08320157.pep
Perfect Score: 1554
Sequence: 1 MASGQGPGPPRQECGEPALP.....LVVLGVVLLGQFVVRRFFKS 211

Scoring table: PAM 150
Gap 11

Searched: 96640 seqs, 11439865 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: a-geneseq27
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20

Statistics: Mean 32.557; Variance 141.390; scale 0.230

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description	Pred. No.
1	1554	100.0	211	20	W03669	Bak-2 protein.	8.16e-138
2	1554	100.0	211	13	R77877	Human Cdn-2.	8.16e-138
3	1529	98.4	211	13	R77876	Human Cdn-1.	2.32e-135
4	1529	98.4	211	20	W03668	Bak protein.	2.32e-135
5	1529	98.4	211	17	R81451	Bcl-Y apoptosis-relat	2.32e-135
6	1136	73.1	152	13	R77879	Human Cdn-1(60-211).	6.78e-97
7	1061	68.3	141	13	R77880	Human Cdn-1(71-211).	1.38e-89
8	880	56.6	116	13	R77881	Human Cdn-1(96-211).	5.18e-72
9	275	17.7	232	19	W01020	Apoptosis-blocking pr	9.24e-15
10	271	17.4	190	13	R68884	Chicken lymphoid BCL-	2.11e-14
11	264	17.0	233	13	R68887	Human thymus BCL-XL.	8.96e-14
12	264	17.0	233	20	W05821	Bcl-XL protein.	8.96e-14
13	262	16.9	232	19	W01019	Apoptosis-blocking pr	1.35e-13
14	262	16.9	239	19	W01018	Apoptosis-blocking pr	1.35e-13
15	260	16.7	205	13	R71405	Human bcl-2 beta prot	2.04e-13
16	260	16.7	205	13	R68886	Human thymus BCL-2.	2.04e-13
17	260	16.7	205	13	R70332	Human bcl-2 protein.	2.04e-13
18	260	16.7	239	13	R71404	Human bcl-2 alpha pro	2.04e-13
19	260	16.7	239	9	R47344	Human oncogene bcl-2	2.04e-13
20	260	16.7	239	1	P80987	Sequence of bcl-2-alp	2.04e-13
21	260	16.7	239	13	R70331	Human bcl-2 protein.	2.04e-13
22	260	16.7	239	8	R42312	Bcl-2 oncogene produc	2.04e-13
23	230	14.8	205	1	P80988	Sequence of bcl-2-bet	9.35e-11
24	222	14.3	229	19	W01021	Apoptosis-blocking pr	4.71e-10
25	219	14.1	192	13	R71407	Murine Bax protein.	8.62e-10
26	213	13.7	88	13	R77878	Human Cdn-3.	2.88e-09
27	208	13.4	192	13	R71406	Human Bax protein.	7.82e-09
28	182	11.7	350	12	R68814	Human mcl-1 gene prod	1.34e-06
29	165	10.6	63	13	R68885	Human thymus BCL-X1.	3.61e-05
30	98	6.3	277	14	R76996	Deduced sequence enco	7.02e+00
31	96	6.2	170	13	R68888	Human thymus BCL-XS.	9.81e+00
32	95	6.1	1367	11	R60795	Human IFG-1 receptor.	1.16e+01
33	95	6.1	1367	17	R91429	Human type I insulin-	1.16e+01
34	95	6.1	1367	17	R95244	IGF-1 receptor.	1.16e+01
35	94	6.0	338	14	R73955	Duffy blood group gpD	1.37e+01
36	93	6.0	740	2	R11516	Soybean chlorotic mot	1.61e+01
37	93	6.0	1337	11	R63125	IGF-I receptor 957A v	1.61e+01
38	93	6.0	1337	11	R63124	IGF-I receptor 950X v	1.61e+01
39	93	6.0	1337	11	R63126	IGF-I receptor 1003A	1.61e+01
40	93	6.0	1337	11	R63123	IGF-I receptor 943A v	1.61e+01
41	93	6.0	1367	11	R63122	IGF-I receptor.	1.61e+01
42	89	5.7	277	15	R79904	ACT-4-h-1 receptor se	3.09e+01
43	89	5.7	1157	9	R48678	Insecticidal protoxin	3.09e+01
44	89	5.7	1984	18	R99639	Peripheral nervous sy	3.09e+01
45	89	5.7	1989	18	R92317	Peripheral nervous sy	3.09e+01

ALIGNMENTS

RESULT 1
ID W03669 standard; Protein; 211 AA.
AC W03669;
DT 22-FEB-1997 (first entry)
DE Bak-2 protein.
KW Human; Bak-2; apoptosis; latency; virus replication;
KW Epstein-Barr virus; BHRF1; fusion protein; epitope tag;
KW drug screening; co-precipitation; ELISA; immunoassay; antibody;
KW protein interactive trapping; virucide; antitumour; diagnostic.
OS Homo sapiens.
PN WO9633416-A1.
PD 24-OCT-1996.
PF 19-APR-1996; U05639.
PR 20-APR-1995; US-426529.
PA (LXRB-) LXR BIOTECHNOLOGY INC.
PI Barr PJ, Kiefer MC;
DR WPI; 96-485886/48.
DR N-PSDB; T42139.
PT Screening for anti-viral agents - by detecting the ability of an
PT agent to disrupt the interaction of a Bak protein and a viral
PT protein
PS Disclosure; Fig 2; 24pp; English.
CC This Bak-2 protein sequence represents a bcl-1 homologue which
CC interacts with Epstein-Barr virus (EBV) early lytic cycle BHRF1
CC protein, and is capable of modulating apoptosis. The protein may
CC be used in complete or partial form, or as an epitope tag fusion
CC protein, in a new virucide drug screening method, which involves
CC combination of Bak-2 protein and a viral protein (e.g. EBV BHRF1),
CC exposure to a test compound, and monitoring for disruption of the
CC interaction, e.g. by co-precipitation, protein interactive trapping
CC or ELISA. Interaction of Bak-2 and viral proteins allows viral
CC replication or latency in the absence of apoptosis. Compounds which
CC inhibit the interaction may be used as virucide, antitumour or
CC diagnostic agents.
SQ Sequence 211 AA;

Query Match 100.0%; Score 1554; DB 20; Length 211;
Best Local Similarity 100.0%; Pred. No. 8.16e-138;
Matches 211; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Db      1 masgqgpgpprqecgepalpsaseeqvaqdteevfrsyvfyhhqqeqeaegaaapadpem 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      1 MASGQGPGPPRQECGEPALPSASEEQVAQDTEEVFRSYVFYHHQQEQEAEGAAAPADPEM 60
```

```
Db      61 vtlplqpsstmgqvgrqlaiigddinrrydsefqtmlqhlqptaenayeyftkiasslfe 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      61 VTLPLQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENAYEYFTKIASSLFE 120

Db     121 sginwgrvvallgfsyrlalhiyqrgltgflgqvtrfvvdfmlhhciarwiaqrggwvaa 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     121 SGINWGRVVALLGFSYRLALHIYQRGLTGFLGQVTRFVVDFMLHHCIARWIAQRGGWVAA 180

Db     181 lnlgngpilnvlvvlgvvllgqfvvrrffks 211
           |||||||||||||||||||||||||||||||
Qy     181 LNLGNGPILNVLVVLGVVLLGQFVVRRFFKS 211


RESULT    2
ID   R77877 standard; Protein; 211 AA.
AC   R77877;
DT   21-NOV-1995  (first entry)
DE   Human Cdn-2.
KW   Cdn-2; apoptosis modulator; adoptive immunotherapy; therapy; HIV;
KW   autoimmune disease; reperfusion injury; hepatitis, osteoporosis;
KW   shock; lymphoma; eczema.
OS   Homo sapiens.
PN   WO9515084-A.
PD   08-JUN-1995.
PF   30-NOV-1994; U13930.
PR   30-NOV-1993; US-160067.
PR   07-OCT-1994; US-320157.
PA   (LXRB-) LXR BIOTECHNOLOGY INC.
PI   Barr PJ,  Kiefer MC;
DR   WPI; 95-215106/28.
DR   N-PSDB; Q95493.
PT   New nucleic acid sequences encoding Cdn apoptosis modulators  - and
PT   related vectors, transformed cells, proteins and antibodies, useful
PT   or diagnosis and treatment e.g. of HIV infection, reperfusion injury
PT   etc.
PS   Disclosure; Fig.5D-E; 66pp; English.
CC   Cdn-2 cDNA was isold. from a human placental genomic library
CC   using a 950 bp fragment of Cdn-1 cDNA.  Expression of Cdn-2
CC   in mouse progenitor B-cell FL5.12 cells decreased IL-3-induced
CC   apoptosis.  The Cdn-2 protein displayed 97% amino acid identity
CC   with Cdn-1 (R77876).
SQ   Sequence   211 AA;

  Query Match          100.0%;  Score 1554;  DB 13;  Length 211;
  Best Local Similarity 100.0%;  Pred. No. 8.16e-138;
  Matches   211;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      1 masgqgpgpprqecgepalpsaseeqvaqdteevfrsyvfyhhqqeqeaegaaapadpem 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      1 MASGQGPGPPRQECGEPALPSASEEQVAQDTEEVFRSYVFYHHQQEQEAEGAAAPADPEM 60

Db     61 vtlplqpsstmgqvgrqlaiigddinrrydsefqtmlqhlqptaenayeyftkiasslfe 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     61 VTLPLQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENAYEYFTKIASSLFE 120

Db    121 sginwgrvvallgfsyrlalhiyqrgltgflgqvtrfvvdfmlhhciarwiaqrggwvaa 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    121 SGINWGRVVALLGFSYRLALHIYQRGLTGFLGQVTRFVVDFMLHHCIARWIAQRGGWVAA 180

Db    181 lnlgngpilnvlvvlgvvllgqfvvrrffks 211
          |||||||||||||||||||||||||||||||
Qy    181 LNLGNGPILNVLVVLGVVLLGQFVVRRFFKS 211


RESULT    3
ID   R77876 standard; Protein; 211 AA.
AC   R77876;
DT   21-NOV-1995  (first entry)
DE   Human Cdn-1.
KW   Cdn-1; apoptosis modulator; adoptive immunotherapy; therapy; HIV;
KW   autoimmune disease; reperfusion injury; hepatitis, osteoporosis;
KW   shock; lymphoma; eczema.
OS   Homo sapiens.
PN   WO9515084-A.
PD   08-JUN-1995.
PF   30-NOV-1994; U13930.
PR   30-NOV-1993; US-160067.
PR   07-OCT-1994; US-320157.
PA   (LXRB-) LXR BIOTECHNOLOGY INC.
PI   Barr PJ,  Kiefer MC;
DR   WPI; 95-215106/28.
DR   N-PSDB; Q95492.
PT   New nucleic acid sequences encoding Cdn apoptosis modulators  - and
PT   related vectors, transformed cells, proteins and antibodies, useful
PT   or diagnosis and treatment e.g. of HIV infection, reperfusion injury
PT   etc.
PS   Disclosure; Fig.3A-B; 66pp; English.
CC   Cdn-1 cDNA was isolated from a human heart cDNA library using a
CC   previously isolated clone as probe.  Recombinant Cdn-1 was produced
CC   in Sf9 and human colon adenocarcinoma HT29 cells.  Expression of
CC   Cdn-1 in WI-L2 lymphoblastoid cells resulted in increased cell
CC   survival in response to anti-Fas-mediated apoptosis.
SQ   Sequence   211 AA;

  Query Match          98.4%;  Score 1529;  DB 13;  Length 211;
  Best Local Similarity 97.2%;  Pred. No. 2.32e-135;
  Matches   205;  Conservative    5;  Mismatches   1;  Indels   0;  Gaps   0;

Db      1 masgqgpgpprqecgepalpsaseeqvaqdteevfrsyvfyrhqqeqeaegvaapadpem 60
          ||||||||||||||||||||||||||||||||||||||||:||||||||| ||||||||
Qy      1 MASGQGPGPPRQECGEPALPSASEEQVAQDTEEVFRSYVFYHHQQEQEAEGAAAPADPEM 60

Db     61 vtlplqpsstmgqvgrqlaiigddinrrydsefqtmlqhlqptaenayeyftkiatslfe 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||
Qy     61 VTLPLQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENAYEYFTKIASSLFE 120

Db    121 sginwgrvvallgfgyrlalhvyqhgltgflgqvtrfvvdfmlhhciarwiaqrggwvaa 180
          |||||||||||||:|||||:||:||||||||||||||||||||||||||||||||||||||
Qy    121 SGINWGRVVALLGFSYRLALHIYQRGLTGFLGQVTRFVVDFMLHHCIARWIAQRGGWVAA 180

Db    181 lnlgngpilnvlvvlgvvllgqfvvrrffks 211
          |||||||||||||||||||||||||||||||
Qy    181 LNLGNGPILNVLVVLGVVLLGQFVVRRFFKS 211


RESULT    4
ID   W03668 standard; Protein; 211 AA.
AC   W03668;
DT   22-FEB-1997  (first entry)
DE   Bak protein.
KW   Human; Bak; apoptosis; latency; virus replication;
KW   Epstein-Barr virus; BHRF1; fusion protein; epitope tag;
KW   drug screening; co-precipitation; ELISA; immunoassay; antibody;
KW   protein interactive trapping; virucide; antitumour; diagnostic.
OS   Homo sapiens.
PN   WO9633416-A1.
PD   24-OCT-1996.
PF   19-APR-1996; U05639.
PR   20-APR-1995; US-426529.
PA   (LXRB-) LXR BIOTECHNOLOGY INC.
PI   Barr PJ,  Kiefer MC;
DR   WPI; 96-485886/48.
DR   N-PSDB; T42138.
PT   Screening for anti-viral agents - by detecting the ability of an
PT   agent to disrupt the interaction of a Bak protein and a viral
PT   protein
PS   Disclosure; Fig 1; 24pp; English.
CC   This Bak protein sequence represents a bcl-1 homologue which
CC   interacts with Epstein-Barr virus (EBV) early lytic cycle BHRF1
CC   protein, and is capable of modulating apoptosis.  The protein may
CC   be used in complete or partial form, or as an epitope tag fusion
CC   protein, in a new virucide drug screening method, which involves
CC   combination of Bak protein and a viral protein (e.g. EBV BHRF1),
CC   exposure to a test compound, and monitoring for disruption of the
CC   interaction, e.g. by co-precipitation, protein interactive trapping
CC   or ELISA.  Interaction of Bak and viral proteins allows viral
```

```
CC   replication or latency in the absence of apoptosis.  Compounds which
CC   inhibit the interaction may be used as virucide, antitumour or
CC   diagnostic agents.
SQ   Sequence   211 AA;

  Query Match           98.4%;  Score 1529;  DB 20;  Length 211;
  Best Local Similarity  97.2%;  Pred. No. 2.32e-135;
  Matches   205;  Conservative    5;  Mismatches   1;  Indels   0;  Gaps   0;

Db      1 masgqgpgpprqecgepalpsaseeqvaqdteevfrsyvfyrhqqeqeaegvaapadpem 60
          ||||||||||||||||||||||||||||||||||||||||||:|||||||| ||||||||
Qy      1 MASGQGPGPPRQECGEPALPSASEEQVAQDTEEVFRSYVFYHHQQEQEAEGAAAPADPEM 60

Db     61 vtlplqpsstmgqvgrqlaiigddinrrydsefqtmlqhlqptaenayeyftkiatslfe 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||
Qy     61 VTLPLQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENAYEYFTKIASSLFE 120

Db    121 sginwgrvvallgfgyrlalhvyqhgltgflgqvtrfvvdfmlhhciarwiaqrggwvaa 180
          |||||||||||||:||||||:||:||||||||||||||||||||||||||||||||||||
Qy    121 SGINWGRVVALLGFSYRLALHIYQRGLTGFLGQVTRFVVDFMLHHCIARWIAQRGGWVAA 180

Db    181 lnlgngpilnvlvvlgvvllgqfvvrrffks 211
          |||||||||||||||||||||||||||||||
Qy    181 LNLGNGPILNVLVVLGVVLLGQFVVRRFFKS 211


RESULT    5
ID   R81451 standard; Protein; 211 AA.
AC   R81451;
DT   02-JUL-1996  (first entry)
DE   Bcl-Y apoptosis-related protein.
KW   Bcl-Y; apoptosis; cell proliferation; cell death; diagnosis;
KW   therapy.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   Domain          188..205
FT   /label= C-terminal_domain
FT   /note= "putative membrane localisation sequence"
PN   WO9605232-A1.
PD   22-FEB-1996.
PF   09-AUG-1995; U10103.
PR   09-AUG-1994; US-287427.
PR   11-OCT-1994; US-321071.
PA   (IMMU-) IMMUNOGEN INC.
PI   Chittenden TD;
DR   WPI; 96-139648/14.
DR   N-PSDB; T17375.
PT   New isolated human Bcl-Y protein - used to develop prods. for
PT   treating disorders characterised by inappropriate cell proliferation
PT   or cell death
PS   Claim 3; Fig 4; 100pp; English.
CC   Bcl-Y protein (R81451) is a member of the Bcl-2 family and can
CC   induce apoptosis in cells and function as a negative regulator of
CC   Bcl-2 function.  Bcl-Y mRNA was detected in all human tumour cell
CC   lines examined and is also widely expressed in primary human
CC   tissues.  It can be obtd. by expression of a full-length cDNA
CC   clone (T17375) in pref. mammalian host cells.  Bcl-Y can be used to
CC   develop prods. for treating disorders associated with inappropriate
CC   cell proliferation or cell death, and to raise antibodies used for
CC   the diagnosis or monitoring of such disorders.
SQ   Sequence   211 AA;

  Query Match           98.4%;  Score 1529;  DB 17;  Length 211;
  Best Local Similarity  97.2%;  Pred. No. 2.32e-135;
  Matches   205;  Conservative    5;  Mismatches   1;  Indels   0;  Gaps   0;

Db      1 masgqgpgpprqecgepalpsaseeqvaqdteevfrsyvfyrhqqeqeaegvaapadpem 60
          ||||||||||||||||||||||||||||||||||||||||||:|||||||| ||||||||
Qy      1 MASGQGPGPPRQECGEPALPSASEEQVAQDTEEVFRSYVFYHHQQEQEAEGAAAPADPEM 60

Db     61 vtlplqpsstmgqvgrqlaiigddinrrydsefqtmlqhlqptaenayeyftkiatslfe 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||
Qy     61 VTLPLQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENAYEYFTKIASSLFE 120

Db    121 sginwgrvvallgfgyrlalhvyqhgltgflgqvtrfvvdfmlhhciarwiaqrggwvaa 180
          |||||||||||||:||||||:||:||||||||||||||||||||||||||||||||||||
Qy    121 SGINWGRVVALLGFSYRLALHIYQRGLTGFLGQVTRFVVDFMLHHCIARWIAQRGGWVAA 180

Db    181 lnlgngpilnvlvvlgvvllgqfvvrrffks 211
          |||||||||||||||||||||||||||||||
Qy    181 LNLGNGPILNVLVVLGVVLLGQFVVRRFFKS 211


RESULT    6
ID   R77879 standard; Protein; 152 AA.
AC   R77879;
DT   21-NOV-1995  (first entry)
DE   Human Cdn-1(60-211).
KW   Cdn-1; apoptosis modulator; adoptive immunotherapy; therapy; HIV;
KW   autoimmune disease; reperfusion injury; hepatitis, osteoporosis;
KW   shock; lymphoma; eczema.
OS   Homo sapiens.
PN   WO9515084-A.
PD   08-JUN-1995.
PF   30-NOV-1994; U13930.
PR   30-NOV-1993; US-160067.
PR   07-OCT-1994; US-320157.
PA   (LXRB-) LXR BIOTECHNOLOGY INC.
PI   Barr PJ,  Kiefer MC;
DR   WPI; 95-215106/28.
PT   New nucleic acid sequences encoding Cdn apoptosis modulators  - and
PT   related vectors, transformed cells, proteins and antibodies, useful
PT   or diagnosis and treatment e.g. of HIV infection, reperfusion injury
PT   etc.
PS   Disclosure; Fig.11; 66pp; English.
CC   Expression of Cdn-1 in WI-L2 lymphoblastoid cells resulted in
CC   increased cell survival in response to anti-Fas-mediated apoptosis.
CC   Deletion of the N-terminal 59 amino acids of Cdn-1 only slightly
CC   decreased this activity, suggesting that small, truncated Cdn-1
CC   molecules may be potent therapeutics.
SQ   Sequence   152 AA;

  Query Match           73.1%;  Score 1136;  DB 13;  Length 152;
  Best Local Similarity  97.4%;  Pred. No. 6.78e-97;
  Matches   148;  Conservative    4;  Mismatches   0;  Indels   0;  Gaps   0;

Db      1 mvtlplqpsstmgqvgrqlaiigddinrrydsefqtmlqhlqptaenayeyftkiatslf 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||
Qy     60 MVTLPLQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENAYEYFTKIASSLF 119

Db     61 esginwgrvvallgfgyrlalhvyqhgltgflgqvtrfvvdfmlhhciarwiaqrggwva 120
          |||||||||||||:||||||:||:||||||||||||||||||||||||||||||||||||
Qy    120 ESGINWGRVVALLGFSYRLALHIYQRGLTGFLGQVTRFVVDFMLHHCIARWIAQRGGWVA 179

Db    121 alnlgngpilnvlvvlgvvllgqfvvrrffks 152
          |||||||||||||||||||||||||||||||
Qy    180 ALNLGNGPILNVLVVLGVVLLGQFVVRRFFKS 211


RESULT    7
ID   R77880 standard; Protein; 141 AA.
AC   R77880;
DT   21-NOV-1995  (first entry)
DE   Human Cdn-1(71-211).
KW   Cdn-1; apoptosis modulator; adoptive immunotherapy; therapy; HIV;
KW   autoimmune disease; reperfusion injury; hepatitis, osteoporosis;
KW   shock; lymphoma; eczema.
OS   Homo sapiens.
PN   WO9515084-A.
PD   08-JUN-1995.
PF   30-NOV-1994; U13930.
PR   30-NOV-1993; US-160067.
PR   07-OCT-1994; US-320157.
PA   (LXRB-) LXR BIOTECHNOLOGY INC.
PI   Barr PJ,  Kiefer MC;
DR   WPI; 95-215106/28.
```

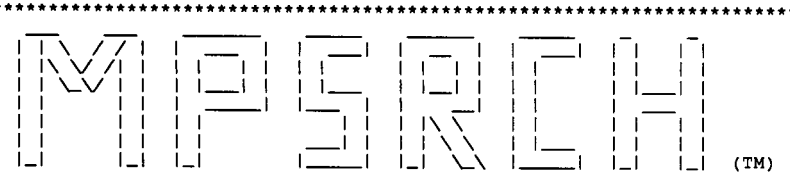
```
PT   New nucleic acid sequences encoding Cdn apoptosis modulators  - and
PT   related vectors, transformed cells, proteins and antibodies, useful
PT   or diagnosis and treatment e.g. of HIV infection, reperfusion injury
PT   etc.
PS   Disclosure; Fig.11; 66pp; English.
CC   Expression of Cdn-1 in WI-L2 lymphoblastoid cells resulted in
CC   increased cell survival in response to anti-Fas-mediated apoptosis.
CC   Deletion of the N-terminal 70 amino acids of Cdn-1 improved this
CC   activity, suggesting that small, truncated Cdn-1 molecules may be
CC   potent therapeutics.
SQ   Sequence   141 AA;

  Query Match           68.3%;  Score 1061;  DB 13;  Length 141;
  Best Local Similarity  97.2%;  Pred. No. 1.38e-89;
  Matches   137;  Conservative    4;  Mismatches   0;  Indels   0;  Gaps   0;

Db      1 mgqvgrqlaiigddinrrydsefqtmlqhlqptaenayeyftkiatslfesginwgrvva 60
          ||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||
Qy     71 MGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENAYEYFTKIASSLFESGINWGRVVA 130

Db     61 llgfgyrlalhvyqhgltgflgqvtrfvvdfmlhhciarwiaqrggwvaalnlgngpiln 120
          ||||:|||||||:||:|||||||||||||||||||||||||||||||||||||||||||||
Qy    131 LLGFSYRLALHIYQRGLTGFLGQVTRFVVDFMLHHCIARWIAQRGGWVAALNLGNGPILN 190

Db    121 vlvvlgvvllgqfvvrrffks 141
          |||||||||||||||||||||
Qy    191 VLVVLGVVLLGQFVVRRFFKS 211


RESULT    8
ID   R77881 standard; Protein; 116 AA.
AC   R77881;
DT   21-NOV-1995  (first entry)
DE   Human Cdn-1(96-211).
KW   Cdn-1; apoptosis modulator; adoptive immunotherapy; therapy; HIV;
KW   autoimmune disease; reperfusion injury; hepatitis, osteoporosis;
KW   shock; lymphoma; eczema.
OS   Homo sapiens.
PN   WO9515084-A.
PD   08-JUN-1995.
PF   30-NOV-1994; U13930.
PR   30-NOV-1993; US-160067.
PR   07-OCT-1994; US-320157.
PA   (LXRB-) LXR BIOTECHNOLOGY INC.
PI   Barr PJ,  Kiefer MC;
DR   WPI; 95-215106/28.
PT   New nucleic acid sequences encoding Cdn apoptosis modulators  - and
PT   related vectors, transformed cells, proteins and antibodies, useful
PT   or diagnosis and treatment e.g. of HIV infection, reperfusion injury
PT   etc.
PS   Disclosure; Fig.11; 66pp; English.
CC   Expression of Cdn-1 in WI-L2 lymphoblastoid cells resulted in
CC   increased cell survival in response to anti-Fas-mediated apoptosis.
CC   Truncated Cdn-1 derivatives given in R77879-81 were used to
CC   test the effects of deleting the N-terminal sequences of Cdn-1
CC   on this activity.
SQ   Sequence   116 AA;

  Query Match           56.6%;  Score 880;  DB 13;  Length 116;
  Best Local Similarity  96.6%;  Pred. No. 5.18e-72;
  Matches   112;  Conservative    4;  Mismatches   0;  Indels   0;  Gaps   0;

Db      1 mlqhlqptaenayeyftkiatslfesginwgrvvallgfgyrlalhvyqhgltgflgqvt 60
          ||||||||||||||||||||:||||||||||||||||:||||||:||:||||||||||
Qy     96 MLQHLQPTAENAYEYFTKIASSLFESGINWGRVVALLGFSYRLALHIYQRGLTGFLGQVT 155

Db     61 rfvvdfmlhhciarwiaqrggwvaalnlgngpilnvlvvlgvvllgqfvvrrffks 116
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    156 RFVVDFMLHHCIARWIAQRGGWVAALNLGNGPILNVLVVLGVVLLGQFVVRRFFKS 211


RESULT    9
ID   W01020 standard; Protein; 232 AA.
AC   W01020;
DT   18-DEC-1996  (first entry)
DE   Apoptosis-blocking protein Bcl-2 mutant 80-6 (del80-86).
KW   Apoptosis-regulating protein; Bcl-2; oncogene;
KW   adenovirus E1B 19K protein; cell death; cancer; tumour;
KW   immune disorder; diagnosis; therapy; Bip1A; Bip13; Bip5; NIp1;
KW   Nip2; Nip3.
OS   Synthetic.
PN   EP-733706-A2.
PD   25-SEP-1996.
PF   21-MAR-1996; 104542.
PR   21-MAR-1995; US-408095.
PA   (UYSL-) UNIV ST LOUIS.
PI   Chinnadurai G;
DR   WPI; 96-427055/43.
PT   Nucleic acids encoding apoptosis regulating proteins - useful for
PT   diagnosing and treating immune disorders, malignancies, etc.
PS   Example 8; Page 34-35; 60pp; English.
CC   The 80-6 mutant (W01020) of the bcl-2 oncogene product (W01018)
CC   lacks amino acids 80-86 of the native protein.  This and other
CC   Bcl-2 mutants (see also W01019-21) were used in a two hybrid assay
CC   to examine the interactions between Bcl-2 and novel apoptosis-
CC   regulating proteins Nip1, Nip2 and Nip3 (W00997-99).  2 Motifs
CC   (W01003-04) on Bcl-2 were identified that are essential for
CC   interaction with the Nip proteins.  These motifs show homology
CC   to motifs (W01005-06) identified on the adenovirus E1B 19K
CC   apoptosis-blocking protein (W01010).
SQ   Sequence   232 AA;

  Query Match           17.7%;  Score 275;  DB 19;  Length 232;
  Best Local Similarity  29.7%;  Pred. No. 9.24e-15;
  Matches    47;  Conservative   41;  Mismatches  62;  Indels   8;  Gaps   6;

Db     48 ffssqpghtphpaasr-dpvartsplqtpaapgspvppvvhltlrqagddfsrryrrdfa 106
          |:  |  : :  ||:  ||   | ||| ::: |  |    :|::   |||::|||  :|
Qy     40 FYHHQQEQEAEGAAAPADPEMVTLPLQPSSTMG-QVGR--QLAI--IGDDINRRYDSEFQ 94

Db    107 emssqlhltpftargcfatvveelfrdgvnwgrivaffefggvmcvesvnremsplvdni 166
           |  :|: |:  |   |: :   ||  |:||||:||:: |:  : :    | :: ::  :
Qy     95 TMLQHLQPTAENAYEYFTKIASSLFESGINWGRVVALLGFSYRLALHIYQRGLTGFLGQV 154

Db    167 alwmteylnrh-lhtwiqdnggwdafvel-ygpsmrpl 202
          :  : ::: :| :  || : ||| | ::|  || :  |
Qy    155 TRFVVDFMLHHCIARWIAQRGGWVAALNLGNGPILNVL 192


RESULT   10
ID   R68884 standard; Protein; 190 AA.
AC   R68884;
DT   10-AUG-1995  (first entry)
DE   Chicken lymphoid BCL-X.
KW   Chicken; bird; fowl; BCL-X; apoptosis; cell death; cancer;
KW   neurodegenerative disease; autoimmune disease; Parkinson's disease;
KW   amylotrophic lateral sclerosis; multiple sclerosis; oncogene.
OS   Gallus domesticus.
PN   WO9500642-A.
PD   05-JAN-1995.
PF   22-JUN-1994; U07089.
PR   22-JUN-1993; US-081448.
PA   (ARCH-) ARCH DEV CORP.
PA   (UNMI ) UNIV MICHIGAN.
PI   Boise LH,  Nunez G,  Thompson CB;
DR   WPI; 95-052079/07.
DR   N-PSDB; Q81696.
PT   New poly-nucleotide encoding new poly-peptide(s) that modify
PT   apoptosis - and related vectors, recombinant cells and
PT   antibodies, useful in assay and for control of cell death in e.g.
PT   neuronal cells, lymphocytes and cancers
PS   Claim 4; Page 87; 127pp; English.
CC   This protein may be expressed recombinantly, particularly with pcmV
CC   plasmids as vectors for expression in mammalian cell cultures.
CC   The protein has particular application in cancer cells (failure of
CC   programmed cell death (PCD)) or neurodegenerative and autoimmune diseases
```

```
CC   (premature PCD), e.g. Parkinson's disease, amylotrophic lateral
CC   sclerosis and multiple sclerosis.
SQ   Sequence   190 AA;

  Query Match           17.4%;  Score 271;  DB 13;  Length 190;
  Best Local Similarity  28.1%;  Pred. No. 2.11e-14;
  Matches    34;  Conservative   36;  Mismatches  49;  Indels   2;  Gaps   2;

Db     70 slevheivrasdvrqalrdagdefelryrrafsdltsqlhitpgtayqsfeqvvnelfhd 129
          :| ::     ::| : |   ||::: ||   |  :  :|: |:  ||: |  : : ||
Qy     62 TLPLQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENAYEYFTKIASSLFES 121

Db    130 gvnwgrivaffsfggalcvesvdkemrvlvgrivswmttyltdh-ldpwiqenggwvrta 188
          |:||||:||:::|:  | :   :: :  ::|::   :  ::  | :  || : |||| :|
Qy    122 GINWGRVVALLGFSYRLALHIYQRGLTGFLGQVTRFVVDFMLHHCIARWIAQRGGWV-AA 180

Db    189 l 189
          |
Qy    181 L 181


RESULT   11
ID   R68887 standard; Protein; 233 AA.
AC   R68887;
DT   10-AUG-1995  (first entry)
DE   Human thymus BCL-XL.
KW   BCL-XL; apoptosis; cell death; cancer; neurodegenerative disease;
KW   autoimmune disease; Parkinson disease; amylotrophic lateral sclerosis;
KW   multiple sclerosis.
OS   Homo sapiens.
PN   WO9500642-A.
PD   05-JAN-1995.
PF   22-JUN-1994; U07089.
PR   22-JUN-1993; US-081448.
PA   (ARCH-) ARCH DEV CORP.
PA   (UNMI ) UNIV MICHIGAN.
PI   Boise LH,  Nunez G,  Thompson CB;
DR   WPI; 95-052079/07.
DR   N-PSDB; Q81698.
PT   New poly-nucleotide encoding new poly-peptide(s) that modify
PT   apoptosis - and related vectors, recombinant cells and
PT   antibodies, useful in assay and for control of cell death in e.g.
PT   neuronal cells, lymphocytes and cancers
PS   Claim 3; Page 94; 127pp; English.
CC   This protein may be expressed recombinantly, particularly with pcmV
CC   plasmids as vectors for expression in mammalian cell cultures.  The
CC   protein has particular application in cancer cells (failure of
CC   programmed cell death (PCD)) or neurodegenerative and autoimmune
CC   diseases (premature PCD), e.g. Parkinson's disease, amylotrophic
CC   lateral sclerosis and multiple sclerosis.
SQ   Sequence   233 AA;

  Query Match           17.0%;  Score 264;  DB 13;  Length 233;
  Best Local Similarity  28.0%;  Pred. No. 8.96e-14;
  Matches    33;  Conservative   37;  Mismatches  46;  Indels   2;  Gaps   2;

Db     83 maavkqalreagdefelryrrafsdltsqlhitpgtayqsfeqvvnelfrdgvnwgriva 142
          |: | : |   ||::: ||   |  :  :|: |:  ||: |  : : ||  |:||||:||
Qy     71 MGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENAYEYFTKIASSLFESGINWGRVVA 130

Db    143 ffsfggalcvesvdkemqvlvsriaawmatylndh-lepwiqenggwdtfvelygnna 199
          :::|:  | :   :: :  :::::  :  ::  | :  || : ||| : ::| || :
Qy    131 LLGFSYRLALHIYQRGLTGFLGQVTRFVVDFMLHHCIARWIAQRGGWVAALNL-GNGP 187


RESULT   12
ID   W05821 standard; Protein; 233 AA.
AC   W05821;
DT   30-MAR-1997  (first entry)
DE   Bcl-XL protein.
KW   Human; bcl-XL; T-lymphocyte; cell death; BH1 domain; BH2 domain;
KW   Bcl-2 homology domain; gene therapy; HIV; AIDS; antisense;
KW   immune disorder; autoimmune disease; graft rejection;
KW   graft-versus-host disease; apoptosis; adoptive immunotherapy.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   Domain          129..148
FT   /note= "BH1 domain"
FT   Domain          180..191
FT   /note= "BH2 domain"
PN   WO9634956-A1.
PD   07-NOV-1996.
PF   02-MAY-1996; U06203.
PR   04-MAY-1995; US-435518.
PR   07-JUN-1995; US-481739.
PA   (ARCH-) ARCH DEV CORP.
PA   (USNA ) US SEC OF NAVY.
PI   June CH,  Thompson CB;
DR   WPI; 96-506159/50.
DR   N-PSDB; T40079.
PT   Inducing or preventing death of T cells by bcl-XL protein regulation
PT   - used to increase survival of HIV infected cells or to
PT   down:regulate immune responses in immune diseases
PS   Disclosure; Page 52-53; 76pp; English.
CC   This is the sequence of a human bcl-XL protein, which protects
CC   T-lymphocytes against cell death.  A splice variant form, bcl-XS,
CC   lacks a stretch of 63 amino acids, and is a dominant negative
CC   regulator of bcl-XL function.  The gene may be modified to
CC   facilitate interaction with costimulatory Bax protein and inhibit
CC   interaction with antagonistic Bad protein, by modification of the
CC   Bcl-2 homology domains BH1 and/or BH2.  The bcl-XL gene may be
CC   introduced into T-cells in vivo or ex vivo via gene transfer using
CC   a vector for HIV infection gene therapy, to augment intracellular
CC   bcl-XL protein levels and protect from cell death.  A corresponding
CC   antisense oligonucleotide or expression vector may be used in gene
CC   therapy of e.g. autoimmune disease, graft rejection or graft-
CC   versus-host disease, to induce cell death (e.g. apoptosis) and
CC   down-regulate the immune response in a T-lymphocyte population.
SQ   Sequence   233 AA;

  Query Match           17.0%;  Score 264;  DB 20;  Length 233;
  Best Local Similarity  28.0%;  Pred. No. 8.96e-14;
  Matches    33;  Conservative   37;  Mismatches  46;  Indels   2;  Gaps   2;

Db     83 maavkqalreagdefelryrrafsdltsqlhitpgtayqsfeqvvnelfrdgvnwgriva 142
          |: | : |   ||::: ||   |  :  :|: |:  ||: |  : : ||  |:||||:||
Qy     71 MGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENAYEYFTKIASSLFESGINWGRVVA 130

Db    143 ffsfggalcvesvdkemqvlvsriaawmatylndh-lepwiqenggwdtfvelygnna 199
          :::|:  | :   :: :  :::::  :  ::  | :  || : ||| : ::| || :
Qy    131 LLGFSYRLALHIYQRGLTGFLGQVTRFVVDFMLHHCIARWIAQRGGWVAALNL-GNGP 187


RESULT   13
ID   W01019 standard; Protein; 232 AA.
AC   W01019;
DT   18-DEC-1996  (first entry)
DE   Apoptosis-blocking protein Bcl-2 mutant 42-8 (del42-48).
KW   Apoptosis-regulating protein; Bcl-2; oncogene;
KW   adenovirus E1B 19K protein; cell death; cancer; tumour;
KW   immune disorder; diagnosis; therapy; Bip1A; Bip13; Bip5; NIp1;
KW   Nip2; Nip3.
OS   Synthetic.
PN   EP-733706-A2.
PD   25-SEP-1996.
PF   21-MAR-1996; 104542.
PR   21-MAR-1995; US-408095.
PA   (UYSL-) UNIV ST LOUIS.
PI   Chinnadurai G;
DR   WPI; 96-427055/43.
PT   Nucleic acids encoding apoptosis regulating proteins - useful for
PT   diagnosing and treating immune disorders, malignancies, etc.
PS   Example 8; Page 33-34; 60pp; English.
CC   The 42-8 mutant (W01019) of the bcl-2 oncogene product (W01018)
CC   lacks amino acids 42-48 of the native protein.  This and other
CC   Bcl-2 mutants (see also W01020-21) were used in a two hybrid assay
```

```
CC   to examine the interactions between Bcl-2 and novel apoptosis-
CC   regulating proteins Nip1, Nip2 and Nip3 (W00997-99).  The Nip
CC   proteins were unable to interact with mutant 42-8.  The site of
CC   deletion in this mutant corresponds to a motif (see also W01003)
CC   on Bcl-2 essential for interaction with Nip proteins.  A second
CC   binding motif (W01004) of Bcl-2 was also identified, and both
CC   show homology to motifs (W01005-06) found on the 19K protein
CC   (W01010) of adenovirus E1B.
SQ   Sequence   232 AA;

  Query Match           16.9%;  Score 262;  DB 19;  Length 232;
  Best Local Similarity  29.3%;  Pred. No. 1.35e-13;
  Matches   41;  Conservative   37;  Mismatches  57;  Indels   5;  Gaps   4;

Db     67 tpaapgaaagpalspvppvvhlt--lrqagddfsrryrrdfaemssqlhltpftargcfa 124
          :|| |   : | | | : : ::   |   |||::|||  :|  |  :|: |:  |   |:
Qy     54 APADPEMVTLP-LQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENAYEYFT 112

Db    125 tvveelfrdgvnwgrivaffefggvmcvesvnremsplvdnialwmteylnrh-lhtwiq 183
           :   ||  |:||||:||:: |:  : :    | :: ::  ::  : ::: :| :  ||
Qy    113 KIASSLFESGINWGRVVALLGFSYRLALHIYQRGLTGFLGQVTRFVVDFMLHHCIARWIA 172

Db    184 dnggwdafvel-ygpsmrpl 202
          : ||| | ::|  || :  |
Qy    173 QRGGWVAALNLGNGPILNVL 192


RESULT   14
ID   W01018 standard; Protein; 239 AA.
AC   W01018;
DT   18-DEC-1996  (first entry)
DE   Apoptosis-blocking protein Bcl-2.
KW   Apoptosis-regulating protein; Bcl-2; oncogene;
KW   adenovirus E1B 19K protein; cell death; cancer; tumour;
KW   immune disorder; diagnosis; therapy; Bip1A; Bip13; Bip5; NIp1;
KW   Nip2; Nip3.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   Binding_site    43..51
FT   /label= Binding_motif
FT   /note= "interacts with Bip proteins"
FT   Binding_site    106..112
FT   /label= Binding_motif
FT   /note= "interacts with Bip proteins"
PN   EP-733706-A2.
PD   25-SEP-1996.
PF   21-MAR-1996; 104542.
PR   21-MAR-1995; US-408095.
PA   (UYSL-) UNIV ST LOUIS.
PI   Chinnadurai G;
DR   WPI; 96-427055/43.
PT   Nucleic acids encoding apoptosis regulating proteins - useful for
PT   diagnosing and treating immune disorders, malignancies, etc.
PS   Example 8; Page 32-33; 60pp; English.
CC   The bcl-2 oncogene product (W01018) enhances the survival of
CC   haematopoietic B and T cells by blocking apoptosis induced by
CC   diverse agents.  Its activity is similar to that of the 19K
CC   protein (W01010) of adenovirus E1B.  3 Novel proteins, Bip1A,
CC   Bip13 and Bip5 (W01000-02), that specifically interact with
CC   Bcl-2, have been identified.  Mutational analysis (see also
CC   W01019-21) shows the apoptosis-regulating Nip proteins (see also
CC   W00997-99) associate with Bcl-2 at specific sites (see also
CC   W01003-04) that show homology to motifs (W01005-06) on 19K.
SQ   Sequence   239 AA;

  Query Match           16.9%;  Score 262;  DB 19;  Length 239;
  Best Local Similarity  29.3%;  Pred. No. 1.35e-13;
  Matches   41;  Conservative   37;  Mismatches  57;  Indels   5;  Gaps   4;

Db     74 tpaapgaaagpalspvppvvhlt--lrqagddfsrryrrdfaemssqlhltpftargcfa 131
          :|| |   : | | | : : ::   |   |||::|||  :|  |  :|: |:  |   |:
Qy     54 APADPEMVTLP-LQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENAYEYFT 112

Db    132 tvveelfrdgvnwgrivaffefggvmcvesvnremsplvdnialwmteylnrh-lhtwiq 190
           :   ||  |:||||:||:: |:  : :    | :: ::  ::  : ::: :| :  ||
Qy    113 KIASSLFESGINWGRVVALLGFSYRLALHIYQRGLTGFLGQVTRFVVDFMLHHCIARWIA 172

Db    191 dnggwdafvel-ygpsmrpl 209
          : ||| | ::|  || :  |
Qy    173 QRGGWVAALNLGNGPILNVL 192


RESULT   15
ID   R71405 standard; protein; 205 AA.
AC   R71405;
DT   30-OCT-1995  (first entry)
DE   Human bcl-2 beta protein.
KW   Human; bcl-2; alpha; beta; proto-oncogene; hematopoietic cell line;
KW   apoptosis; membrane-associated cytoplasmic protein; B cell; T cell;
KW   proliferation; cell cycle progression; Bax; apoptotic cell death;
KW   apoptosis; cytokine;  death repressor; BH1; BH2; cancer therapy;
KW   hyperplasia;  immunodeficiency disease; AIDS; neurodegeneration;
KW   ischaemic cell death.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   Domain          136..155
FT   /label= BH1_domain
FT   /note= "Represents Bax binding site"
FT   Domain          187..192
FT   /label= BH2_domain
FT   /note= "Represents Bax binding site"
PN   WO9505750-A.
PD   02-MAR-1995.
PF   24-AUG-1994; U09701.
PR   26-AUG-1993; US-112208.
PR   25-MAY-1994; US-248819.
PA   (UNIW ) UNIV WASHINGTON.
PI   Korsmeyer SJ;
DR   WPI; 95-106605/14.
PT   Methods for producing and identifying mutant bcl-2 proteins -
PT   that lack death repressor activity and/or lacks binding to Bax.
PS   Disclosure; Page 40; 133pp; English.
CC   The sequences given in R71404-05 represent the human bcl-2 alpha and
CC   beta proteins respectively.  bcl-2 is encoded by a proto-oncogene and
CC   is capable of inhibiting apoptosis in many hematopoietic cell systems.
CC   bcl-2 is a 26 kD membrane-associated cytoplasmic protein and is thought
CC   to function by enhancing the survival of hematopoietic cells of B and T
CC   origins rather than directly promoting proliferation of these cell
CC   types.  bcl-2 has not been shown to directly promote cell cycle
CC   progression nor does it necessarily alter the dose reponse to limiting
CC   concentrations of IL-3.  bcl-2 has been shown to form heterodimers with
CC   a 21 kD protein, Bax.  Overexpressed Bax accelerates apoptotic cell death
CC   induced by cytokine deprivation in an IL-3 dependent cell line, and it
CC   also acts to counter the death repressor activity of bcl-2.  Therefore,
CC   the ratio between bcl-2 and Bax determines cell survival or death
CC   following an apoptotic stimulus.  The invention gives a mutant form of
CC   bcl-2 in which there is at least one amino acid substitution or deletion
CC   in the BH1 or BH2 domains.  This makes the mutant protein substantially
CC   incapable of binding Bax and/or incapable of death repressor activity.
CC   Down regulation of bcl-2 is useful in cancer therapy, controlling
CC   hyperplasias and eliminating self-reactive clones in autoimmunity by
CC   favouring death effector molecules.  Up regulating bcl-2is beneficial in
CC   treatment and diagnosis of immunodeficiency diseases, including AIDS and
CC   neurodegenerative and ischaemic cell death.
SQ   Sequence   205 AA;

  Query Match           16.7%;  Score 260;  DB 13;  Length 205;
  Best Local Similarity  29.7%;  Pred. No. 2.04e-13;
  Matches   38;  Conservative   37;  Mismatches  49;  Indels   4;  Gaps   3;

Db     74 tpaapgaaagpalspvppvvhla--lrqagddfsrryrgdfaemssqlhltpftargrfa 131
          :|| |   : | | | : : :::   |   |||::|||  ::|  |  :|: |:  |   |:
Qy     54 APADPEMVTLP-LQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENAYEYFT 112

Db    132 tvveelfrdgvnwgrivaffefggvmcvesvnremsplvdnialwmteylnrh-lhtwiq 190
           :   ||  |:||||:||:: |:  : :    | :: ::  ::  : ::: :| :  ||
```

```
Qy    113 KIASSLFESGINWGRVVALLGFSYRLALHIYQRGLTGFLGQVTRFVVDFMLHHCIARWIA 172

Db    191 dnggwvga 198
          : ||||:|
Qy    173 QRGGWVAA 180
```

Search completed: Wed Aug 20 11:05:16 1997
Job time : 22 secs.



MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Wed Aug 20 11:06:16 1997; MasPar time 6.19 Seconds
723.119 Million cell updates/sec

Tabular output not generated.

Title: >US-08-320-157-9
Description: (1-211) from US08320157.pep
Perfect Score: 1554
Sequence: 1 MASGQGPGPPRQECGEPALP.....LVVLGVVLLGQFVVRRFFKS 211

Scoring table: PAM 150
Gap 11

Searched: 59021 seqs, 21210388 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: swiss-prot34
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11

Statistics: Mean 47.039; Variance 90.944; scale 0.517

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description	Pred. No.
1	271	17.4	190	1	BCLX_CHICK	APOPTOSIS REGULATOR B	1.55e-29
2	268	17.2	233	1	BCLX_RAT	APOPTOSIS REGULATOR B	5.65e-29
3	266	17.1	233	1	BCL2_CHICK	APOPTOSIS REGULATOR B	1.34e-28
4	264	17.0	233	1	BCLX_HUMAN	APOPTOSIS REGULATOR B	3.16e-28
5	256	16.5	205	1	BC2B_HUMAN	PROTEIN BCL-2-BETA.	9.72e-27
6	256	16.5	239	1	BC2A_HUMAN	PROTEIN BCL-2-ALPHA.	9.72e-27
7	251	16.2	199	1	BC2B_MOUSE	PROTEIN BCL-2-BETA.	8.18e-26
8	251	16.2	236	1	BC2A_MOUSE	PROTEIN BCL-2 ALPHA.	8.18e-26
9	250	16.1	236	1	BCL2_RAT	PROTEIN BCL-2 ALPHA.	1.25e-25
10	221	14.2	192	1	BAXA_MOUSE	APOPTOSIS REGULATOR B	2.37e-20
11	208	13.4	192	1	BAXA_HUMAN	APOPTOSIS REGULATOR B	4.87e-18
12	203	13.1	218	1	BAXB_HUMAN	APOPTOSIS REGULATOR B	3.69e-17
13	178	11.5	143	1	BAXD_HUMAN	BAX PROTEIN, CYTOPLAS	7.38e-13
14	179	11.5	350	6	MCL1_HUMAN	INDUCED MYELOID LEUKE	5.01e-13
15	173	11.1	172	5	HSA1_MOUSE	HEMOPOIETIC-SPECIFIC	5.09e-12
16	145	9.3	179	3	EAR_ASFE4	APOPTOSIS REGULATOR B	1.73e-07
17	142	9.1	179	3	EAR_ASFB7	APOPTOSIS REGULATOR B	5.05e-07
18	139	8.9	179	3	EAR_ASFM2	APOPTOSIS REGULATOR B	1.46e-06
19	115	7.4	626	5	HTPG_BACSU	HEAT SHOCK PROTEIN HT	4.77e-03
20	107	6.9	255	2	COX3_CAEEL	CYTOCHROME C OXIDASE	5.81e-02
21	105	6.8	567	5	HXT3_YEAST	LOW-AFFINITY GLUCOSE	1.07e-01
22	105	6.8	576	5	HXT4_YEAST	LOW-AFFINITY GLUCOSE	1.07e-01
23	104	6.7	133	4	GOP_BPP4	GOP PROTEIN.	1.44e-01
24	103	6.6	470	9	THRC_PSEAE	THREONINE SYNTHASE (E	1.94e-01
25	102	6.6	569	5	HXT8_YEAST	HEXOSE TRANSPORTER HX	2.61e-01
26	103	6.6	817	8	RRPO_TBSVC	PROBABLE RNA-DIRECTED	1.94e-01
27	99	6.4	527	11	YAD7_SCHPO	HYPOTHETICAL 59.6 KD	6.28e-01
28	100	6.4	567	5	HXT9_YEAST	HEXOSE TRANSPORTER HX	4.70e-01
29	100	6.4	567	5	HXTA_YEAST	HEXOSE TRANSPORTER HX	4.70e-01
30	98	6.3	680	8	PTFB_MYCGE	PTS SYSTEM, FRUCTOSE-	8.37e-01
31	96	6.2	562	3	CSTA_ECOLI	CARBON STARVATION PRO	1.48e+00
32	97	6.2	570	5	HXT6_YEAST	HIGH AFFINITY HEXOSE	1.11e+00
33	97	6.2	570	5	HXT7_YEAST	HIGH AFFINITY HEXOSE	1.11e+00
34	97	6.2	833	3	CW41_YEAST	PROTEIN CWH41 PRECURS	1.11e+00
35	96	6.2	1323	5	HST6_CANAL	ATP-DEPENDENT PERMEAS	1.48e+00
36	95	6.1	462	11	YNX6_CAEEL	HYPOTHETICAL 51.6 KD	1.96e+00
37	95	6.1	566	5	KHT2_KLULA	HEXOSE TRANSPORTER 2.	1.96e+00
38	95	6.1	822	11	YJC9_YEAST	HYPOTHETICAL 95.4 KD	1.96e+00
39	95	6.1	885	1	AR56_SCHPO	ARG11 PROTEIN PRECURS	1.96e+00
40	95	6.1	1367	5	IG1R_HUMAN	INSULIN-LIKE GROWTH F	1.96e+00
41	94	6.0	817	8	RRPO_CRV	PROBABLE RNA-DIRECTED	2.59e+00
42	94	6.0	817	8	RRPO_CNV	PROBABLE RNA-DIRECTED	2.59e+00
43	92	5.9	196	6	MLEF_HUMAN	MYOSIN LIGHT CHAIN 1,	4.49e+00
44	92	5.9	570	5	HXT1_YEAST	LOW-AFFINITY GLUCOSE	4.49e+00
45	92	5.9	919	7	PMA1_SCHPO	PLASMA MEMBRANE ATPAS	4.49e+00

ALIGNMENTS

```
RESULT    1
ID   BCLX_CHICK     STANDARD;      PRT;   190 AA.
AC   Q07816;
DT   01-FEB-1995 (REL. 31, CREATED)
DT   01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)
DT   01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)
DE   APOPTOSIS REGULATOR BCL-X.
GN   BCL-X.
OS   GALLUS GALLUS (CHICKEN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; AVES; NEOGNATHAE;
OC   GALLIFORMES.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 93364977.
RA   BOISE L.H., GONZALEZ-GARCIA M., POSTEMA C.E., DING L., LINDSTEN T.,
RA   TURKA L.A., MAO X., NUNEZ G., THOMPSON C.B.;
RL   CELL 74:597-608(1993).
CC   -!- FUNCTION: DOMINANT REGULATOR OF APOPTOTIC CELL DEATH.
CC   -!- TISSUE SPECIFICITY: HIGHEST EXPRESSION IN ORGANS WITH
CC       LYMPHOID DEVELOPMENT.
CC   -!- SIMILARITY: BELONGS TO THE BCL-2 FAMILY.
DR   EMBL; Z23110; G510899; -.
DR   PIR; A47537; A47537.
DR   PROSITE; PS01080; BCL2.
KW   APOPTOSIS.
SQ   SEQUENCE   190 AA;  21467 MW;  7874E430 CRC32;

  Query Match           17.4%;  Score 271;  DB 1;  Length 190;
  Best Local Similarity  28.1%;  Pred. No. 1.55e-29;
  Matches    34;  Conservative   36;  Mismatches  49;  Indels   2;  Gaps   2;

Db     70 slevheivrasdvrqalrdagdefelryrrafsdltsqlhitpgtayqsfeqvvnelfhd 129
          :| ::     ::| : |   ||::: ||   |  :  :|: |:  ||: |  : : ||
Qy     62 TLPLQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENAYEYFTKIASSLFES 121

Db    130 gvnwgrivaffsfggalcvesvdkemrvlvgrivswmttyltdh-ldpwiqenggwvrta 188
          |:||||:||:::|:  | :   :: :  ::|::   :  ::  | :  || : |||| :|
Qy    122 GINWGRVVALLGFSYRLALHIYQRGLTGFLGQVTRFVVDFMLHHCIARWIAQRGGWV-AA 180

Db    189 l 189
          |
Qy    181 L 181


RESULT    2
ID   BCLX_RAT       STANDARD;      PRT;   233 AA.
```

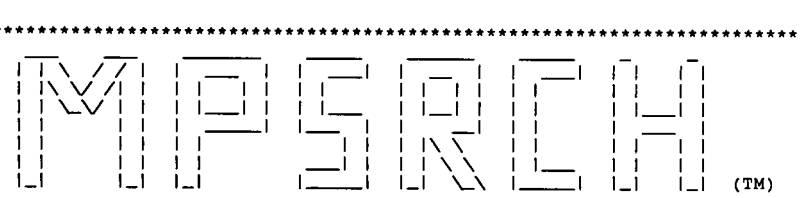
```
AC   P53563;
DT   01-OCT-1996 (REL. 34, CREATED)
DT   01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   APOPTOSIS REGULATOR BCL-X.
GN   BCLX.
OS   RATTUS NORVEGICUS (RAT).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; RODENTIA.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=BRAIN;
RA   MICHAELIDIS T.M.;
RL   SUBMITTED (DEC-1994) TO EMBL/GENBANK/DDBJ DATA BANKS.
CC   -!- FUNCTION: DOMINANT REGULATOR OF APOPTOTIC CELL DEATH.
CC   -!- ALTERNATIVE PRODUCTS: TWO ISOFORMS, BCX-X(L) (SHOWN HERE) AND
CC       BCL-X(S), ARE DERIVED BY ALTERNATIVE SPLICING.
CC   -!- SIMILARITY: BELONGS TO THE BCL-2 FAMILY.
DR   EMBL; X82537; G607177; -.
DR   EMBL; X82537; G607178; -.
KW   APOPTOSIS; ALTERNATIVE SPLICING.
FT   VARSPLIC    126    188       MISSING (IN BCL-X(S)).
SQ   SEQUENCE   233 AA;  26130 MW;  E0589815 CRC32;

  Query Match           17.2%;  Score 268;  DB 1;  Length 233;
  Best Local Similarity  28.0%;  Pred. No. 5.65e-29;
  Matches   33;  Conservative   37;  Mismatches  46;  Indels   2;  Gaps   2;

Db     83 maavkqalreagdefelryrrafsdltsqlhitpgtayqsfeqvvnelfrdgvnwgriva 142
          |: | : |   ||::: ||   |  :  :|: |:  ||: |  : : ||  |:||||:||
Qy     71 MGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENAYEYFTKIASSLFESGINWGRVVA 130

Db    143 ffsfggalcvesvdkemqvlvsriaswmatylndh-lepwiqenggwdtfvdlygnna 199
          :::|:  | :   :: :  :::::  :  ::  | :  || : ||| : ::| || :
Qy    131 LLGFSYRLALHIYQRGLTGFLGQVTRFVVDFMLHHCIARWIAQRGGWVAALNL-GNGP 187


RESULT    3
ID   BCL2_CHICK     STANDARD;      PRT;   233 AA.
AC   Q00709;
DT   01-APR-1993 (REL. 25, CREATED)
DT   01-APR-1993 (REL. 25, LAST SEQUENCE UPDATE)
DT   01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE   APOPTOSIS REGULATOR BCL-2.
GN   BCL-2.
OS   GALLUS GALLUS (CHICKEN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; AVES; NEOGNATHAE;
OC   GALLIFORMES.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 92375724.
RA   EGUCHI Y., EWERT D.L., TSUJIMOTO Y.;
RL   NUCLEIC ACIDS RES. 20:4187-4192(1992).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 92379084.
RA   CAZALS-HATEM D.L., LOUIE D.C., TANAKA S., REED J.C.;
RL   BIOCHIM. BIOPHYS. ACTA 1132:109-113(1992).
CC   -!- FUNCTION: PROLONGS THE SURVIVAL OF HEMATOPOIETIC CELLS IN THE
CC       ABSENCE OF REQUIRED GROWTH FACTORS AND ALSO IN THE PRESENCE OF
CC       VARIOUS STIMULI INDUCING CELLULAR DEATH. BLOCKS APOPTOSIS. MIGHT
CC       FUNCTION IN AN ANTIOXIDANT PATHWAY TO PREVENT APOPTOSIS AT SITES
CC       OF FREE RADICAL GENERATION SUCH AS MITOCHONDRIA.
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL INNER MEMBRANE
CC       (BY SIMILARITY).
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN A VARIETY OF TISSUES.
CC   -!- SIMILARITY: BELONGS TO THE BCL-2 FAMILY.
DR   EMBL; D11382; G222794; -.
DR   EMBL; D11381; G222794; JOINED.
DR   EMBL; Z11961; G62970; -.
DR   PIR; A37332; A37332.
DR   PIR; S24390; S24390.
DR   PROSITE; PS01080; BCL2.
KW   APOPTOSIS; TRANSMEMBRANE; MITOCHONDRION.
FT   TRANSMEM    208    228       POTENTIAL.
FT   CONFLICT     64     64       E -> S (IN REF. 2).
FT   CONFLICT     67     82       GSAAASEVPPAEGLRP -> ARLLLVRCPRLRGCA
FT                                (IN REF. 2).
FT   CONFLICT    121    121       H -> T (IN REF. 2).
FT   CONFLICT    139    139       G -> V (IN REF. 2).
SQ   SEQUENCE   233 AA;  25687 MW;  3376502C CRC32;

  Query Match           17.1%;  Score 266;  DB 1;  Length 233;
  Best Local Similarity  28.5%;  Pred. No. 1.34e-28;
  Matches   41;  Conservative   42;  Mismatches  56;  Indels   5;  Gaps   5;

Db     60 hhrpeppgsaaasevppae-glrpapp-g-vhlalrqagdefsrryqrdfaqmsgqlhlt 116
          ::  |: |:||:::       |:|::  | |   |   ||:::|||: :|  |  :|: |
Qy     44 QQEQEAEGAAAPADPEMVTLPLQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPT 103

Db    117 pftahgrfvavveelfrdgvnwgrivaffefggvmcvesvnremsplvdniatwmteyln 176
          :  |   |  :   ||  |:||||:||:: |:  : :    | :: ::  ::  : :::
Qy    104 AENAYEYFTKIASSLFESGINWGRVVALLGFSYRLALHIYQRGLTGFLGQVTRFVVDFML 163

Db    177 rh-lhnwiqdnggwdafvelygns 199
          :| :  || : ||| | ::| ||:
Qy    164 HHCIARWIAQRGGWVAALNL-GNG 186


RESULT    4
ID   BCLX_HUMAN     STANDARD;      PRT;   233 AA.
AC   Q07817;
DT   01-FEB-1995 (REL. 31, CREATED)
DT   01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   APOPTOSIS REGULATOR BCL-X.
GN   BCLX.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 93364977.
RA   BOISE L.H., GONZALEZ-GARCIA M., POSTEMA C.E., DING L.,
RA   LINDSTEN T., TURKA L.A., MAO X., NUNEZ G., THOMPSON C.B.;
RL   CELL 74:597-608(1993).
CC   -!- FUNCTION: DOMINANT REGULATOR OF APOPTOTIC CELL DEATH.
CC   -!- ALTERNATIVE PRODUCTS: TWO ISOFORMS, BCX-X(L) (SHOWN HERE) AND
CC       BCL-X(S), ARE DERIVED BY ALTERNATIVE SPLICING.
CC   -!- TISSUE SPECIFICITY: BCL-X(S) IS EXPRESSED AT HIGH LEVELS IN CELLS
CC       THAT UNDERGO A HIGH RATE OF TURNOVER, SUCH AS DEVELOPING
CC       LYMPHOCYTES. IN CONTRAST, BCL-X(L) IS FOUND IN TISSUES CONTAINING
CC       LONG-LIVED POSTMITOTIC CELLS, SUCH AS ADULT BRAIN.
CC   -!- SIMILARITY: BELONGS TO THE BCL-2 FAMILY.
DR   EMBL; Z23116; G623237; -.
DR   EMBL; Z23115; G510901; -.
DR   MIM; 600039; -.
DR   PROSITE; PS01080; BCL2.
KW   APOPTOSIS; ALTERNATIVE SPLICING.
FT   VARSPLIC    126    188       MISSING (IN BCL-X(S)).
FT   CONFLICT     70     70       G -> A (IN G510901).
SQ   SEQUENCE   233 AA;  26049 MW;  57C67491 CRC32;

  Query Match           17.0%;  Score 264;  DB 1;  Length 233;
  Best Local Similarity  28.0%;  Pred. No. 3.16e-28;
  Matches   33;  Conservative   37;  Mismatches  46;  Indels   2;  Gaps   2;

Db     83 maavkqalreagdefelryrrafsdltsqlhitpgtayqsfeqvvnelfrdgvnwgriva 142
          |: | : |   ||::: ||   |  :  :|: |:  ||: |  : : ||  |:||||:||
Qy     71 MGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENAYEYFTKIASSLFESGINWGRVVA 130

Db    143 ffsfggalcvesvdkemqvlvsriaawmatylndh-lepwiqenggwdtfvelygnna 199
          :::|:  | :   :: :  :::::  :  ::  | :  || : ||| : ::| || :
Qy    131 LLGFSYRLALHIYQRGLTGFLGQVTRFVVDFMLHHCIARWIAQRGGWVAALNL-GNGP 187
```

```
RESULT    5
ID   BC2B_HUMAN     STANDARD;      PRT;   205 AA.
AC   P10416;
DT   01-MAR-1989 (REL. 10, CREATED)
DT   01-APR-1993 (REL. 25, LAST SEQUENCE UPDATE)
DT   01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE   PROTEIN BCL-2-BETA.
GN   BCL2.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 86259760.
RA   TSUJIMOTO Y., CROCE C.M.;
RL   PROC. NATL. ACAD. SCI. U.S.A. 83:5214-5218(1986).
RN   [2]
RP   REVISIONS TO 96 AND 110.
RX   MEDLINE; 92375724.
RA   EGUCHI Y., EWERT D.L., TSUJIMOTO Y.;
RL   NUCLEIC ACIDS RES. 20:4187-4192(1992).
RN   [3]
RP   SUBCELLULAR LOCATION.
RX   MEDLINE; 91066924.
RA   HOCKENBERY D., NUNEZ G., MILLIMAN C., SCHREIBER R.D., KORSMEYER S.J.;
RL   NATURE 348:334-336(1990).
CC   -!- FUNCTION: PROLONGS THE SURVIVAL OF HEMATOPOIETIC CELLS IN THE
CC       ABSENCE OF REQUIRED GROWTH FACTORS AND ALSO IN THE PRESENCE OF
CC       VARIOUS STIMULI INDUCING CELLULAR DEATH. BLOCKS APOPTOSIS. MIGHT
CC       FUNCTION IN AN ANTIOXIDANT PATHWAY TO PREVENT APOPTOSIS AT SITES
CC       OF FREE RADICAL GENERATION SUCH AS MITOCHONDRIA.
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL INNER MEMBRANE.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN A VARIETY OF TISSUES.
CC   -!- DISEASE: INVOLVED IN FOLLICULAR LYMPHOMA (FL) (ALSO KNOWN AS TYPE
CC       II CHRONIC LYMPHATIC LEUKEMIA) BY A CHROMOSOMAL TRANSLOCATION
CC       T(14;18)(Q32;Q21) WHICH INVOLVES BCL2 AND IMMUNOGLOBULIN GENE
CC       REGIONS.
CC   -!- ALTERNATIVE PRODUCTS: TWO FORMS OF BCL-2: ALPHA, AND BETA, ARE
CC       PRODUCED BY ALTERNATIVE SPLICING OF THE SAME GENE. THEY ONLY
CC       DIFFER AT THEIR C-TERMINAL ENDS.
CC   -!- SIMILARITY: BELONGS TO THE BCL-2 FAMILY.
DR   EMBL; M13995; G179369; ALT_SEQ.
DR   PIR; B29409; TVHUB1.
DR   PIR; D37332; D37332.
DR   MIM; 151430; -.
DR   PROSITE; PS01080; BCL2.
KW   PROTO-ONCOGENE; APOPTOSIS; ALTERNATIVE SPLICING; MEMBRANE;
KW   MITOCHONDRION; CHROMOSOMAL TRANSLOCATION.
SQ   SEQUENCE   205 AA;  22311 MW;  ED321E5E CRC32;

  Query Match           16.5%;  Score 256;  DB 1;  Length 205;
  Best Local Similarity  29.7%;  Pred. No. 9.72e-27;
  Matches    38;  Conservative   35;  Mismatches  51;  Indels   4;  Gaps   3;

Db     74 tpaapgaaagpalspvppvvhlt--lrqagddfsrryrrdfaemssqlhltpftargrfa 131
          :|| |   : | | | : : ::   |   |||::|||  :|  |  :|: |:  |   |:
Qy     54 APADPEMVTLP-LQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENAYEYFT 112

Db    132 tvveelfrdgvnwgrivaffefggvmcvesvnremsplvdnialwmteylnrh-lhtwiq 190
           :   ||  |:||||:||:: |:  : :    | :: ::  ::  : ::: :| :  ||
Qy    113 KIASSLFESGINWGRVVALLGFSYRLALHIYQRGLTGFLGQVTRFVVDFMLHHCIARWIA 172

Db    191 dnggwvga 198
          : ||||:|
Qy    173 QRGGWVAA 180


RESULT    6
ID   BC2A_HUMAN     STANDARD;      PRT;   239 AA.
AC   P10415;
DT   01-MAR-1989 (REL. 10, CREATED)
DT   01-APR-1993 (REL. 25, LAST SEQUENCE UPDATE)
DT   01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE   PROTEIN BCL-2-ALPHA.
GN   BCL2.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 86259760.
RA   TSUJIMOTO Y., CROCE C.M.;
RL   PROC. NATL. ACAD. SCI. U.S.A. 83:5214-5218(1986).
RN   [2]
RP   REVISIONS TO 96; 110 AND 237.
RX   MEDLINE; 92375724.
RA   EGUCHI Y., EWERT D.L., TSUJIMOTO Y.;
RL   NUCLEIC ACIDS RES. 20:4187-4192(1992).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 87002488.
RA   CLEARY M.L., SMITH S.D., SKLAR J.;
RL   CELL 47:19-28(1986).
RN   [4]
RP   SUBCELLULAR LOCATION.
RX   MEDLINE; 91066924.
RA   HOCKENBERY D., NUNEZ G., MILLIMAN C., SCHREIBER R.D., KORSMEYER S.J.;
RL   NATURE 348:334-336(1990).
CC   -!- FUNCTION: PROLONGS THE SURVIVAL OF HEMATOPOIETIC CELLS IN THE
CC       ABSENCE OF REQUIRED GROWTH FACTORS AND ALSO IN THE PRESENCE OF
CC       VARIOUS STIMULI INDUCING CELLULAR DEATH. BLOCKS APOPTOSIS. MIGHT
CC       FUNCTION IN AN ANTIOXIDANT PATHWAY TO PREVENT APOPTOSIS AT SITES
CC       OF FREE RADICAL GENERATION SUCH AS MITOCHONDRIA.
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL INNER MEMBRANE.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN A VARIETY OF TISSUES.
CC   -!- DISEASE: INVOLVED IN FOLLICULAR LYMPHOMA (FL) (ALSO KNOWN AS TYPE
CC       II CHRONIC LYMPHATIC LEUKEMIA) BY A CHROMOSOMAL TRANSLOCATION
CC       T(14;18)(Q32;Q21) WHICH INVOLVES BCL2 AND IMMUNOGLOBULIN GENE
CC       REGIONS.
CC   -!- ALTERNATIVE PRODUCTS: TWO FORMS OF BCL-2: ALPHA, AND BETA, ARE
CC       PRODUCED BY ALTERNATIVE SPLICING OF THE SAME GENE. THEY ONLY
CC       DIFFER AT THEIR C-TERMINAL ENDS.
CC   -!- SIMILARITY: BELONGS TO THE BCL-2 FAMILY.
DR   EMBL; M13994; G179367; ALT_SEQ.
DR   EMBL; M14745; G179371; -.
DR   PIR; A29409; TVHUA1.
DR   PIR; A24428; TVHUBC.
DR   PIR; C37332; C37332.
DR   MIM; 151430; -.
DR   PROSITE; PS01080; BCL2.
KW   PROTO-ONCOGENE; APOPTOSIS; ALTERNATIVE SPLICING; TRANSMEMBRANE;
KW   MITOCHONDRION; CHROMOSOMAL TRANSLOCATION.
FT   TRANSMEM    212    233       POTENTIAL.
FT   CONFLICT     59     59       P -> T (IN REF. 3).
FT   CONFLICT    117    117       S -> R (IN REF. 3).
SQ   SEQUENCE   239 AA;  26266 MW;  75084B59 CRC32;

  Query Match           16.5%;  Score 256;  DB 1;  Length 239;
  Best Local Similarity  29.3%;  Pred. No. 9.72e-27;
  Matches    41;  Conservative   37;  Mismatches  57;  Indels   5;  Gaps   4;

Db     74 tpaapgaaagpalspvppvvhlt--lrqagddfsrryrrdfaemssqlhltpftargrfa 131
          :|| |   : | | | : : ::   |   |||::|||  :|  |  :|: |:  |   |:
Qy     54 APADPEMVTLP-LQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENAYEYFT 112

Db    132 tvveelfrdgvnwgrivaffefggvmcvesvnremsplvdnialwmteylnrh-lhtwiq 190
           :   ||  |:||||:||:: |:  : :    | :: ::  ::  : ::: :| :  ||
Qy    113 KIASSLFESGINWGRVVALLGFSYRLALHIYQRGLTGFLGQVTRFVVDFMLHHCIARWIA 172

Db    191 dnggwdafvel-ygpsmrpl 209
          : ||| | ::|  || :  |
Qy    173 QRGGWVAALNLGNGPILNVL 192


RESULT    7
ID   BC2B_MOUSE     STANDARD;      PRT;   199 AA.
```

```
AC   P10418;
DT   01-MAR-1989 (REL. 10, CREATED)
DT   01-MAR-1989 (REL. 10, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   PROTEIN BCL-2-BETA.
GN   BCL2 OR BCL-2.
OS   MUS MUSCULUS (MOUSE).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; RODENTIA.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 87187643.
RA   NEGRINI M., SILINI E., KOZAK C., TSUJIMOTO Y., CROCE C.M.;
RL   CELL 49:455-463(1987).
CC   -!- FUNCTION: PROLONGS THE SURVIVAL OF HEMATOPOIETIC CELLS IN THE
CC       ABSENCE OF REQUIRED GROWTH FACTORS AND ALSO IN THE PRESENCE OF
CC       VARIOUS STIMULI INDUCING CELLULAR DEATH. BLOCKS APOPTOSIS. MIGHT
CC       FUNCTION IN AN ANTIOXIDANT PATHWAY TO PREVENT APOPTOSIS AT SITES
CC       OF FREE RADICAL GENERATION SUCH AS MITOCHONDRIA.
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL INNER MEMBRANE.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN A VARIETY OF TISSUES.
CC   -!- ALTERNATIVE PRODUCTS: TWO FORMS OF BCL-2: ALPHA, AND BETA, ARE
CC       PRODUCED BY ALTERNATIVE SPLICING OF THE SAME GENE. THEY ONLY
CC       DIFFER AT THEIR C-TERMINAL ENDS.
CC   -!- SIMILARITY: BELONGS TO THE BCL-2 FAMILY.
DR   EMBL; M16506; G387110; -.
DR   PIR; B25960; TVMSB1.
DR   PROSITE; PS01080; BCL2.
KW   APOPTOSIS; ALTERNATIVE SPLICING; MEMBRANE; MITOCHONDRION.
SQ   SEQUENCE   199 AA;  22299 MW;  3E60FA47 CRC32;

  Query Match           16.2%;  Score 251;  DB 1;  Length 199;
  Best Local Similarity  31.1%;  Pred. No. 8.18e-26;
  Matches    32;  Conservative   29;  Mismatches  41;  Indels   1;  Gaps   1;

Db     94 lrragddfsrryrrdfaemssqlhltpftargrfatvveelfrdgvnwgrivaffefggv 153
          |   |||::|||  :|  |  :|: |:  |   |: :   ||  |:||||:||:: |:
Qy     78 LAIIGDDINRRYDSEFQTMLQHLQPTAENAYEYFTKIASSLFESGINWGRVVALLGFSYR 137

Db    154 mcvesvnremsplvdnialwmteylnrh-lhtwiqdnggwvga 195
          : :    | :: ::  ::  : ::: :| :  || : ||||:|
Qy    138 LALHIYQRGLTGFLGQVTRFVVDFMLHHCIARWIAQRGGWVAA 180


RESULT    8
ID   BC2A_MOUSE     STANDARD;      PRT;   236 AA.
AC   P10417;
DT   01-MAR-1989 (REL. 10, CREATED)
DT   01-APR-1993 (REL. 25, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   PROTEIN BCL-2 ALPHA.
GN   BCL2 OR BCL-2.
OS   MUS MUSCULUS (MOUSE).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; RODENTIA.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/C; TISSUE=LIVER;
RX   MEDLINE; 87187643.
RA   NEGRINI M., SILINI E., KOZAK C., TSUJIMOTO Y., CROCE C.M.;
RL   CELL 49:455-463(1987).
RN   [2]
RP   REVISIONS TO 221-222.
RX   MEDLINE; 92375724.
RA   EGUCHI Y., EWERT D.L., TSUJIMOTO Y.;
RL   NUCLEIC ACIDS RES. 20:4187-4192(1992).
CC   -!- FUNCTION: PROLONGS THE SURVIVAL OF HEMATOPOIETIC CELLS IN THE
CC       ABSENCE OF REQUIRED GROWTH FACTORS AND ALSO IN THE PRESENCE OF
CC       VARIOUS STIMULI INDUCING CELLULAR DEATH. BLOCKS APOPTOSIS. MIGHT
CC       FUNCTION IN AN ANTIOXIDANT PATHWAY TO PREVENT APOPTOSIS AT SITES
CC       OF FREE RADICAL GENERATION SUCH AS MITOCHONDRIA.
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL INNER MEMBRANE.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN A VARIETY OF TISSUES.
CC   -!- ALTERNATIVE PRODUCTS: TWO FORMS OF BCL-2: ALPHA, AND BETA, ARE
CC       PRODUCED BY ALTERNATIVE SPLICING OF THE SAME GENE. THEY ONLY
CC       DIFFER AT THEIR C-TERMINAL ENDS.
CC   -!- SIMILARITY: BELONGS TO THE BCL-2 FAMILY.
DR   EMBL; L31532; G387109; -.
DR   EMBL; M16506; G387109; JOINED.
DR   PIR; A25960; TVMSA1.
DR   PIR; E37332; E37332.
DR   PROSITE; PS01080; BCL2.
KW   APOPTOSIS; ALTERNATIVE SPLICING; TRANSMEMBRANE; MITOCHONDRION.
FT   TRANSMEM    209    230       POTENTIAL.
SQ   SEQUENCE   236 AA;  26425 MW;  7ADFE975 CRC32;

  Query Match           16.2%;  Score 251;  DB 1;  Length 236;
  Best Local Similarity  30.4%;  Pred. No. 8.18e-26;
  Matches    35;  Conservative   31;  Mismatches  47;  Indels   2;  Gaps   2;

Db     94 lrragddfsrryrrdfaemssqlhltpftargrfatvveelfrdgvnwgrivaffefggv 153
          |   |||::|||  :|  |  :|: |:  |   |: :   ||  |:||||:||:: |:
Qy     78 LAIIGDDINRRYDSEFQTMLQHLQPTAENAYEYFTKIASSLFESGINWGRVVALLGFSYR 137

Db    154 mcvesvnremsplvdnialwmteylnrh-lhtwiqdnggwdafvel-ygpsmrpl 206
          : :    | :: ::  ::  : ::: :| :  || : ||| | ::|  || :  |
Qy    138 LALHIYQRGLTGFLGQVTRFVVDFMLHHCIARWIAQRGGWVAALNLGNGPILNVL 192


RESULT    9
ID   BCL2_RAT       STANDARD;      PRT;   236 AA.
AC   P49950;
DT   01-OCT-1996 (REL. 34, CREATED)
DT   01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   PROTEIN BCL-2 ALPHA.
GN   BCL2 OR BCL-2.
OS   RATTUS NORVEGICUS (RAT).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; RODENTIA.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=BRAIN;
RX   MEDLINE; 94193015.
RA   SATO T., IRIE S., KRAJEWSKI S., REED J.C.;
RL   GENE 140:291-292(1994).
CC   -!- FUNCTION: PROLONGS THE SURVIVAL OF HEMATOPOIETIC CELLS IN THE
CC       ABSENCE OF REQUIRED GROWTH FACTORS AND ALSO IN THE PRESENCE OF
CC       VARIOUS STIMULI INDUCING CELLULAR DEATH. BLOCKS APOPTOSIS. MIGHT
CC       FUNCTION IN AN ANTIOXIDANT PATHWAY TO PREVENT APOPTOSIS AT SITES
CC       OF FREE RADICAL GENERATION SUCH AS MITOCHONDRIA.
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL INNER MEMBRANE.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN A VARIETY OF TISSUES (BY
CC       SIMILARITY).
CC   -!- ALTERNATIVE PRODUCTS: TWO FORMS OF BCL-2: ALPHA, AND BETA, ARE
CC       PRODUCED BY ALTERNATIVE SPLICING OF THE SAME GENE. THEY ONLY
CC       DIFFER AT THEIR C-TERMINAL ENDS.
CC   -!- SIMILARITY: BELONGS TO THE BCL-2 FAMILY.
DR   EMBL; L14680; G408947; -.
KW   APOPTOSIS; ALTERNATIVE SPLICING; TRANSMEMBRANE; MITOCHONDRION.
FT   TRANSMEM    209    230       POTENTIAL.
SQ   SEQUENCE   236 AA;  26550 MW;  336E6B40 CRC32;

  Query Match           16.1%;  Score 250;  DB 1;  Length 236;
  Best Local Similarity  27.2%;  Pred. No. 1.25e-25;
  Matches    40;  Conservative   43;  Mismatches  60;  Indels   4;  Gaps   4;

Db     63 rdtaartsplrp-lvanagpalspvppvvhltlrragddfsrryrrdfaemssqlhltpf 121
          ::: : ::|  | :|: :  : | :  | :  |  |||::|||  :|  |  :|: |:
Qy     47 QEAEGAAAPADPEMVTLPLQPSSTMGQVGR-QLAIIGDDINRRYDSEFQTMLQHLQPTAE 105

Db    122 targrfatvveelfrdgvnwgrivaffefggvmcvgsvnremsplvdnialwmteylnrh 181
            |   |: :   ||  |:||||:||:: |:  : :    | :: ::  ::  : ::: :|
Qy    106 NAYEYFTKIASSLFESGINWGRVVALLGFSYRLALHIYQRGLTGFLGQVTRFVVDFMLHH 165

Db    182 -lhtwiqdnggwdafvel-ygpsmrpl 206
```

```
          :  || : ||| | ::|  || :  |
Qy    166 CIARWIAQRGGWVAALNLGNGPILNVL 192


RESULT   10
ID   BAXA_MOUSE     STANDARD;      PRT;   192 AA.
AC   Q07813;
DT   01-FEB-1995 (REL. 31, CREATED)
DT   01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   APOPTOSIS REGULATOR BAX, MEMBRANE ISOFORM ALPHA.
GN   BAX.
OS   MUS MUSCULUS (MOUSE).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; RODENTIA.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6 X DBA/2 F1;
RX   MEDLINE; 93364978.
RA   OLTVAI Z.N., MILLIMAN C.L., KORSMEYER S.J.;
RL   CELL 74:609-619(1993).
CC   -!- FUNCTION: ACCELERATES PROGRAMMED CELL DEATH BY BINDING TO THE
CC       APOPTOSIS REPRESSOR BCL-2.
CC   -!- ALTERNATIVE PRODUCTS: A 21 KD MEMBRANE PROTEIN ALPHA AND THE TWO
CC       CYTOPLASMIC PROTEINS BETA AND GAMMA ARE GENERATED BY ALTERNATIVE
CC       SPLICING.
CC   -!- SUBUNIT: FORMS HOMODIMERS AND HETERODIMERS TOGETHER WITH BCL-2.
CC   -!- SUBCELLULAR LOCATION: MEMBRANE.
CC   -!- SIMILARITY: BELONGS TO THE BCL-2 FAMILY.
DR   EMBL; L22472; G388192; -.
DR   PROSITE; PS01080; BCL2.
KW   APOPTOSIS; TRANSMEMBRANE; ALTERNATIVE SPLICING.
FT   TRANSMEM    172    192       POTENTIAL.
SQ   SEQUENCE   192 AA;  21394 MW;  BD035304 CRC32;

  Query Match           14.2%;  Score 221;  DB 1;  Length 192;
  Best Local Similarity  25.6%;  Pred. No. 2.37e-20;
  Matches    46;  Conservative   50;  Mismatches  75;  Indels   9;  Gaps   8;

Db     16 seqimktgafllqgfiqdragrmagetpeltleqppqdastk-klseclrrigdelds-- 72
          :|:: ::  |  :   |:  |  |   ||: :  | | :||:  ::  |  |||:::
Qy     31 TEEVFRSYVFYHHQQEQEAEGAAAPADPEM-VTLPLQPSSTMGQVGRQLAIIGDDINRRY 89

Db     73 nmelqrmiadvdtdspre-vff-rvaadmfadgnfnwgrvvalfyfasklvlkalctkvp 130
          : |:| |:  ::  :     :| ::|: :|  | :||||||||: |: :| |       :
Qy     90 DSEFQTMLQHLQPTAENAYEYFTKIASSLFESG-INWGRVVALLGFSYRLALHIYQRGLT 148

Db    131 elirtimgwtldf-lrerllvwiqdqggwegllsyfgtptwqtvtifvagvltasltiwk 189
          ::  :    :|| |:  :  || ::||| : |:    |    | : :: || : : :::
Qy    149 GFLGQVTRFVVDFMLHHCIARWIAQRGGWVAALNLGNGPI-LNVLVVLGVVLLGQFVVRR 207


RESULT   11
ID   BAXA_HUMAN     STANDARD;      PRT;   192 AA.
AC   Q07812;
DT   01-FEB-1995 (REL. 31, CREATED)
DT   01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   APOPTOSIS REGULATOR BAX, MEMBRANE ISOFORM ALPHA.
GN   BAX.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=B-CELL;
RX   MEDLINE; 93364978.
RA   OLTVAI Z.N., MILLIMAN C.L., KORSMEYER S.J.;
RL   CELL 74:609-619(1993).
CC   -!- FUNCTION: ACCELERATES PROGRAMMED CELL DEATH BY BINDING TO THE
CC       APOPTOSIS REPRESSOR BCL-2.
CC   -!- SUBUNIT: FORMS HOMODIMERS AND HETERODIMERS TOGETHER WITH BCL-2.
CC   -!- SUBCELLULAR LOCATION: MEMBRANE.
CC   -!- ALTERNATIVE PRODUCTS: THE MEMBRANE ISOFORM ALPHA AND THE THREE
CC       CYTOPLASMIC ISOFORMS, BETA, GAMMA AND DELTA ARE GENERATED BY
CC       ALTERNATIVE SPLICING.
CC   -!- SIMILARITY: BELONGS TO THE BCL-2 FAMILY.
DR   EMBL; L22473; G388166; -.
DR   PIR; A47538; A47538.
DR   MIM; 600040; -.
DR   PROSITE; PS01080; BCL2.
KW   APOPTOSIS; TRANSMEMBRANE; ALTERNATIVE SPLICING.
FT   TRANSMEM    172    192       POTENTIAL.
SQ   SEQUENCE   192 AA;  21184 MW;  B2E6148A CRC32;

  Query Match           13.4%;  Score 208;  DB 1;  Length 192;
  Best Local Similarity  24.5%;  Pred. No. 4.87e-18;
  Matches    40;  Conservative   45;  Mismatches  71;  Indels   7;  Gaps   6;

Db     32 qdragrmggeapelaldpvpqdastkklseclkrigdelds--nmelqrmiaavdtdspr 89
          |:  |   :   ||:   |:  ::   ::  |  |||:::   : |:| |:  ::  :
Qy     47 QEAEGAAAPADPEMVTLPLQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAEN 106

Db     90 e-vff-rvaadmfsdgnfnwgrvvalfyfasklvlkalctkvpelirtimgwtldf-lre 146
            :| ::|: :|  | :|||||||||: |: :| |       :  ::  :    :|| |:
Qy    107 AYEYFTKIASSLFESG-INWGRVVALLGFSYRLALHIYQRGLTGFLGQVTRFVVDFMLHH 165

Db    147 rllgwiqdqggwdgllsyfgtptwqtvtifvagvltasltiwk 189
           :  || ::||| : |:    |    | : :: || : : :::
Qy    166 CIARWIAQRGGWVAALNLGNGPI-LNVLVVLGVVLLGQFVVRR 207


RESULT   12
ID   BAXB_HUMAN     STANDARD;      PRT;   218 AA.
AC   Q07814;
DT   01-FEB-1995 (REL. 31, CREATED)
DT   01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   APOPTOSIS REGULATOR BAX, CYTOPLASMIC ISOFORM BETA.
GN   BAX.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=B-CELL;
RX   MEDLINE; 93364978.
RA   OLTVAI Z.N., MILLIMAN C.L., KORSMEYER S.J.;
RL   CELL 74:609-619(1993).
CC   -!- FUNCTION: ACCELERATES PROGRAMMED CELL DEATH BY BINDING TO THE
CC       APOPTOSIS REPRESSOR BCL-2.
CC   -!- SUBUNIT: FORMS HOMODIMERS AND HETERODIMERS TOGETHER WITH BCL-2.
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC   -!- ALTERNATIVE PRODUCTS: THE MEMBRANE ISOFORM ALPHA AND THE THREE
CC       CYTOPLASMIC ISOFORMS, BETA, GAMMA AND DELTA ARE GENERATED BY
CC       ALTERNATIVE SPLICING.
CC   -!- SIMILARITY: BELONGS TO THE BCL-2 FAMILY.
DR   EMBL; L22474; G388168; -.
DR   PIR; B47538; B47538.
DR   MIM; 600040; -.
DR   PROSITE; PS01080; BCL2.
KW   APOPTOSIS; ALTERNATIVE SPLICING.
SQ   SEQUENCE   218 AA;  24220 MW;  82B2FF09 CRC32;

  Query Match           13.1%;  Score 203;  DB 1;  Length 218;
  Best Local Similarity  27.0%;  Pred. No. 3.69e-17;
  Matches    37;  Conservative   36;  Mismatches  58;  Indels   6;  Gaps   5;

Db     32 qdragrmggeapelaldpvpqdastkklseclkrigdelds--nmelqrmiaavdtdspr 89
          |:  |   :   ||:   |:  ::   ::  |  |||:::   : |:| |:  ::  :
Qy     47 QEAEGAAAPADPEMVTLPLQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAEN 106

Db     90 e-vff-rvaadmfsdgnfnwgrvvalfyfasklvlkalctkvpelirtimgwtldf-lre 146
            :| ::|: :|  | :|||||||||: |: :| |       :  ::  :    :|| |:
Qy    107 AYEYFTKIASSLFESG-INWGRVVALLGFSYRLALHIYQRGLTGFLGQVTRFVVDFMLHH 165
```

```
Db    147 rllgwiqdqggwvrllk 163
          :  || ::|||| |:
Qy    166 CIARWIAQRGGWVAALN 182


RESULT   13
ID   BAXD_HUMAN     STANDARD;      PRT;   143 AA.
AC   P55269;
DT   01-OCT-1996 (REL. 34, CREATED)
DT   01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   BAX PROTEIN, CYTOPLASMIC ISOFORM DELTA.
GN   BAX.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 95331797.
RA   APTE S.S., MATTEI M.-G., OLSEN B.R.;
RL   GENOMICS 26:592-594(1995).
CC   -!- ALTERNATIVE PRODUCTS: THE MEMBRANE ISOFORM ALPHA AND THE THREE
CC       CYTOPLASMIC ISOFORMS, BETA, GAMMA AND DELTA ARE GENERATED BY
CC       ALTERNATIVE SPLICING.
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC (POTENTIAL).
CC   -!- SIMILARITY: BELONGS TO THE BCL-2 FAMILY.
DR   EMBL; U19599; G841238; -.
DR   MIM; 600040; -.
KW   APOPTOSIS; ALTERNATIVE SPLICING.
SQ   SEQUENCE   143 AA;  15772 MW;  71AA1CBD CRC32;

  Query Match           11.5%;  Score 178;  DB 1;  Length 143;
  Best Local Similarity  27.4%;  Pred. No. 7.38e-13;
  Matches    29;  Conservative   31;  Mismatches  43;  Indels   3;  Gaps   3;

Db     36 tdsprevffrvaadmfsdgnfnwgrvvalfyfasklvlkalctkvpelirtimgwtldf- 94
          :::: | | ::|: :|  | :|||||||||: |: :| |      :  ::  :    :||
Qy    104 AENAYEYFTKIASSLFESG-INWGRVVALLGFSYRLALHIYQRGLTGFLGQVTRFVVDFM 162

Db     95 lrerllgwiqdqggwdgllsyfgtptwqtvtifvagvltasltiwk 140
          |:  :  || ::||| : |:    |    | : :: || : : :::
Qy    163 LHHCIARWIAQRGGWVAALNLGNGPI-LNVLVVLGVVLLGQFVVRR 207


RESULT   14
ID   MCL1_HUMAN     STANDARD;      PRT;   350 AA.
AC   Q07820;
DT   01-FEB-1995 (REL. 31, CREATED)
DT   01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)
DT   01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)
DE   INDUCED MYELOID LEUKEMIA CELL DIFFERENTIATION PROTEIN MCL1.
GN   MCL1.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=MYELOID LEUKEMIA CELLS;
RX   MEDLINE; 93234528.
RA   KOZOPAS K.M., YANG T., BUCHAN H.L., ZHOU P., CRAIG R.W.;
RL   PROC. NATL. ACAD. SCI. U.S.A. 90:3516-3520(1993).
CC   -!- FUNCTION: INVOLVED IN PROGRAMMING OF DIFFERENTIATION AND
CC       CONCOMITANT MAINTENANCE OF VIABILITY BUT NOT OF PROLIFERATION
CC       (PROBABLE).
CC   -!- INDUCTION: EXPRESSION INCREASES EARLY DURING PHORBOL-ESTER INDUCED
CC       DIFFERENTIATION ALONG THE MONOCYTE/MAGROPHAGE PATHWAY IN MYELOID
CC       LEUKEMIA CELL LINES ML-1.
CC   -!- SIMILARITY: BELONGS TO THE BCL2 FAMILY.
DR   EMBL; L08246; -; NOT_ANNOTATED_CDS.
DR   PIR; A47476; A47476.
DR   MIM; 159552; -.
DR   PROSITE; PS01080; BCL2.
KW   APOPTOSIS; TRANSMEMBRANE; DIFFERENTIATION.
FT   UNSURE      227    227       OR A.
FT   TRANSMEM    330    349       POTENTIAL.
SQ   SEQUENCE   350 AA;  37365 MW;  10194B64 CRC32;

  Query Match           11.5%;  Score 179;  DB 6;  Length 350;
  Best Local Similarity  25.7%;  Pred. No. 5.01e-13;
  Matches    26;  Conservative   29;  Mismatches  44;  Indels   2;  Gaps   2;

Db    213 lrrvgdgvqrnhetvfqgmlrkldikneddvkslsrvmihvfsdgvtnwgrivtlisfga 272
          |  :|| : |  :: || ||: |:   |:    ::::   :|  |: ||||:|:|::|:
Qy     78 LAIIGDDINRRYDSEFQTMLQHLQPTAENAYEYFTKIASSLFESGI-NWGRVVALLGFSY 136

Db    273 fvakhlktinqescieplaesitd-vlvrtkrdwlvkqrgw 312
           :| |:      : :  ::  : | :| :    |:  : ||
Qy    137 RLALHIYQRGLTGFLGQVTRFVVDFMLHHCIARWIAQRGGW 177


RESULT   15
ID   HSA1_MOUSE     STANDARD;      PRT;   172 AA.
AC   Q07440;
DT   01-FEB-1995 (REL. 31, CREATED)
DT   01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)
DT   01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)
DE   HEMOPOIETIC-SPECIFIC EARLY RESPONSE PROTEIN (A1 PROTEIN).
GN   A1.
OS   MUS MUSCULUS (MOUSE).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; RODENTIA.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CBA/J; TISSUE=BONE MARROW;
RX   MEDLINE; 93346743.
RA   LIN E.Y., ORLOFSKY A., BERGER M.S., PRYSTOWSKY M.B.;
RL   J. IMMUNOL. 151:1979-1988(1993).
CC   -!- FUNCTION: MAY FUNCTION IN THE RESPONSE OF HEMOPOIETIC CELLS TO
CC       EXTERNAL SIGNALS.
CC   -!- INDUCTION: BY GRANULOCYTE-MACROPHAGE COLONY-STIMULATING FACTOR
CC       AND LPS IN MACROPHAGES.
CC   -!- SUBCELLULAR LOCATION: INTRACELLULAR.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN HEMOPOIETIC TISSUES, INCLUDING
CC       BONE MARROW, SPLEEN AND THYMUS.
CC   -!- SIMILARITY: BELONGS TO THE BCL-2 FAMILY.
DR   EMBL; L16462; G293274; -.
DR   PROSITE; PS01080; BCL2.
KW   APOPTOSIS.
FT   DOMAIN       24     33       ALA/PRO-RICH.
SQ   SEQUENCE   172 AA;  19914 MW;  FFD38D6F CRC32;

  Query Match           11.1%;  Score 173;  DB 5;  Length 172;
  Best Local Similarity  28.1%;  Pred. No. 5.09e-12;
  Matches    32;  Conservative   30;  Mismatches  46;  Indels   6;  Gaps   6;

Db     28 sapsqacrvlqrvafsvqkeveknlksylddfhvesidtariifnqvmekefedgiinwg 87
          |: :|  | |  ::  : :  : :: : |: ::  : : |   |  :    || || |||
Qy     69 STMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQ-PTAENAYEYFTKIASSLFESGI-NWG 126

Db     88 rivtifafggvllkklpqeqialdvcaykqvssfvaefimnn-tgewirqnggw 140
          |:|::::|:   |  |   | :|      ||: || :|:::  : || | |||
Qy    127 RVVALLGFS-YRLA-LHIYQRGLTG-FLGQVTRFVVDFMLHHCIARWIAQRGGW 177


Search completed: Wed Aug 20 11:06:31 1997
Job time : 15 secs.
```



MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Wed Aug 20 11:05:33 1997; MasPar time 8.89 Seconds
685.993 Million cell updates/sec

Tabular output not generated.

Title: >US-08-320-157-9
Description: (1-211) from US08320157.pep
Perfect Score: 1554
Sequence: 1 MASGQGPGPPRQECGEPALP.....LVVLGVVLLGQFVVRRFFKS 211

Scoring table: PAM 150
Gap 11

Searched: 91006 seqs, 28888923 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: pir51
1:ann1 2:ann2 3:ann3 4:ann4 5:unann1 6:unann2 7:unann3
8:unann4 9:unann5 10:unann6 11:unann7 12:unann8
13:unann9 14:unann10 15:unenc 16:unrev

Statistics: Mean 45.185; Variance 108.828; scale 0.415

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description	Pred. No.
1	1554	100.0	211	13	S58875	cdn-2 protein - huma	1.43e-241
2	1529	98.4	211	13	S58873	Bak protein - human	3.45e-237
3	271	17.4	190	13	A47537	apoptosis regulator	1.24e-23
4	267	17.2	216	6	B37332	transforming protein	5.03e-23
5	268	17.2	233	14	S51761	BCL-X protein - rat	3.55e-23
6	268	17.2	233	14	I49056	bcl-x long - mouse	3.55e-23
7	266	17.1	233	6	A37332	transforming protein	7.14e-23
8	264	17.0	233	13	B47537	apoptosis regulator	1.43e-22
9	261	16.8	214	14	I49057	bcl-x transmembrane	4.08e-22
10	260	16.7	205	2	TVHUB1	transforming protein	5.77e-22
11	256	16.5	206	6	D37332	transforming protein	2.31e-21
12	256	16.5	239	2	TVHUA1	transforming protein	2.31e-21
13	251	16.2	199	2	TVMSB1	transforming protein	1.30e-20
14	251	16.2	233	14	I67431	BCL-X-Long - rat	1.30e-20
15	251	16.2	236	2	TVMSA1	transforming protein	1.30e-20
16	251	16.2	237	6	E37332	transforming protein	1.30e-20
17	249	16.0	236	14	I67432	BCL-2 - rat (fragmen	2.60e-20
18	240	15.4	232	6	S24390	transforming protein	5.72e-19
19	220	14.2	192	14	D47538	programmed cell deat	5.02e-16
20	208	13.4	192	13	A47538	bcl-2-associated pro	2.74e-14
21	203	13.1	218	13	B47538	bcl-2-associated pro	1.43e-13
22	187	12.0	133	14	I53295	bax - rat (fragment)	2.61e-11
23	182	11.7	154	14	I58194	gene bcl-2 protein -	1.29e-10
24	178	11.5	143	13	I38921	BAX splice form delt	4.62e-10
25	179	11.5	350	13	A47476	BCL2 homolog MCL1 -	3.36e-10
26	173	11.1	172	14	I49449	hemopoietic-specific	2.24e-09
27	117	7.5	400	10	S35958	beta-glucosidase - S	3.17e-02
28	114	7.3	177	13	S54778	NR-13 protein - quai	7.04e-02
29	107	6.9	255	5	S26032	cytochrome-c oxidase	4.33e-01
30	107	6.9	890	8	S44150	coat protein - straw	4.33e-01
31	105	6.8	560	7	S46724	hexose transport pro	7.19e-01
32	105	6.8	567	7	S31294	hexose transport pro	7.19e-01
33	105	6.8	567	11	S70110	HXT3 protein - yeast	7.19e-01
34	104	6.7	133	4	GPBPP4	gop protein - satell	9.24e-01
35	103	6.6	470	2	SYPSRA	threonine synthase (	1.19e+00
36	102	6.6	569	7	S50771	sugar transport prot	1.52e+00
37	103	6.6	817	1	RRVGCT	RNA-directed RNA pol	1.19e+00
38	101	6.5	2007	3	B43402	myosin heavy chain-B	1.95e+00
39	99	6.4	527	16	S62484	hypothetical protein	3.17e+00
40	100	6.4	567	7	S49600	glucose transport pr	2.49e+00
41	100	6.4	567	7	S50708	sugar transport prot	2.49e+00
42	100	6.4	640	5	S23008	insulin-like growth	2.49e+00
43	98	6.3	680	10	H64206	fructose-permease II	4.04e+00
44	97	6.2	570	11	S70107	HXT7 protein - yeast	5.13e+00
45	97	6.2	570	16	S67932	hexose transporter H	5.13e+00

ALIGNMENTS

RESULT 1
ENTRY S58875 #type complete
TITLE cdn-2 protein - human
ORGANISM #formal_name Homo sapiens #common_name man
DATE 15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 01-Mar-1996
ACCESSIONS S58875
REFERENCE S58874
#authors Kiefer, M.C.; Brauer, M.J.; Powers, V.C.; Wu, J.J.; Umansky, S.R.; Tomei, L.D.; Barr, P.J.
#journal Nature (1995) 374:736-739
#title Modulation of apoptosis by the widely distributed Bcl-2 homologue Bak.
#accession S58875
##status preliminary; nucleic acid sequence not shown; translation not shown
##molecule_type DNA
##residues 1-211 ##label KIE
##cross-references EMBL:U16812
##note the nucleotide sequence was submitted to the EMBL Data Library, November 1994
SUMMARY #length 211 #molecular-weight 23411 #checksum 9485

Query Match 100.0%; Score 1554; DB 13; Length 211;
Best Local Similarity 100.0%; Pred. No. 1.43e-241;
Matches 211; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Db      1 masgqgpgpprqecgepalpsaseeqvaqdteevfrsyvfyhhqqeqeaegaaapadpem 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      1 MASGQGPGPPRQECGEPALPSASEEQVAQDTEEVFRSYVFYHHQQEQEAEGAAAPADPEM 60

Db     61 vtlplqpsstmgqvgrqlaiigddinrrydsefqtmlqhlqptaenayeyftkiasslfe 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     61 VTLPLQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENAYEYFTKIASSLFE 120

Db    121 sginwgrvvallgfsyrlalhiyqrgltgflgqvtrfvvdfmlhhciarwiaqrggwvaa 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    121 SGINWGRVVALLGFSYRLALHIYQRGLTGFLGQVTRFVVDFMLHHCIARWIAQRGGWVAA 180

Db    181 lnlgngpilnvlvvlgvvllgqfvvrrffks 211
          |||||||||||||||||||||||||||||||
Qy    181 LNLGNGPILNVLVVLGVVLLGQFVVRRFFKS 211
```

RESULT 2

```
ENTRY            S58873     #type complete
TITLE            Bak protein - human
ALTERNATE_NAMES  bcl-2 homolog; cdn-1 protein
ORGANISM         #formal_name Homo sapiens #common_name man
DATE             15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change
                   06-Sep-1996
ACCESSIONS       S58873; S58872; S58874
REFERENCE        S58873
   #authors      Chittenden, T.; Harrington, E.A.; O'Connor, R.; Flemington,
                   C.; Lutz, R.J.; Evan, G.I.; Guild, B.C.
   #journal      Nature (1995) 374:733-736
   #title        Induction of apoptosis by the Bcl-2 homologue Bak.
   #accession    S58873
      ##status        preliminary; nucleic acid sequence not shown
      ##molecule_type mRNA
      ##residues      1-211 ##label CHI
      ##cross-references EMBL:U23765
REFERENCE        S58872
   #authors      Farrow, S.N.; White, J.H.M.; Martinou, I.; Raven, T.; Pun,
                   K.T.; Grinham, C.J.; Martinou, J.C.; Brown, R.
   #journal      Nature (1995) 374:731-733
   #title        Cloning of a bcl-2 homologue by interaction with adenovirus
                   E1B 19K.
   #accession    S58872
      ##status        preliminary
      ##molecule_type mRNA
      ##residues      1-211 ##label FAR
      ##cross-references EMBL:X84213
REFERENCE        S58874
   #authors      Kiefer, M.C.; Brauer, M.J.; Powers, V.C.; Wu, J.J.; Umansky,
                   S.R.; Tomei, L.D.; Barr, P.J.
   #journal      Nature (1995) 374:736-739
   #title        Modulation of apoptosis by the widely distributed Bcl-2
                   homologue Bak.
   #accession    S58874
      ##status        preliminary
      ##molecule_type mRNA
      ##residues      1-211 ##label KIE
      ##cross-references EMBL:U16811
GENETICS
   #gene         GDB:BAK-LSB
      ##cross-references GDB:635887
SUMMARY          #length 211  #molecular-weight 23409  #checksum 801

  Query Match           98.4%;  Score 1529;  DB 13;  Length 211;
  Best Local Similarity  97.2%;  Pred. No. 3.45e-237;
  Matches   205;  Conservative    5;  Mismatches   1;  Indels   0;  Gaps   0;

Db       1 masgqgpgpprqecgepalpsaseeqvaqdteevfrsyvfyrhqqeqeaegvaapadpem 60
           ||||||||||||||||||||||||||||||||||||||||:||||||||| ||||||||
Qy       1 MASGQGPGPPRQECGEPALPSASEEQVAQDTEEVFRSYVFYHHQQEQEAEGAAAPADPEM 60

Db      61 vtlplqpsstmgqvgrqlaiigddinrrydsefqtmlqhlqptaenayeyftkiatslfe 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||
Qy      61 VTLPLQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENAYEYFTKIASSLFE 120

Db     121 sginwgrvvallgfgyrlalhvyqhgltgflgqvtrfvvdfmlhhciarwiaqrggwvaa 180
           ||||||||||||:||||||:||:||||||||||||||||||||||||||||||||||||
Qy     121 SGINWGRVVALLGFSYRLALHIYQRGLTGFLGQVTRFVVDFMLHHCIARWIAQRGGWVAA 180

Db     181 lnlgngpilnvlvvlgvvllgqfvvrrffks 211
           |||||||||||||||||||||||||||||||
Qy     181 LNLGNGPILNVLVVLGVVLLGQFVVRRFFKS 211


RESULT    3
ENTRY            A47537     #type complete
TITLE            apoptosis regulator bcl-x - chicken
ORGANISM         #formal_name Gallus gallus #common_name chicken
DATE             03-May-1994 #sequence_revision 03-May-1994 #text_change
                   03-May-1994
ACCESSIONS       A47537
REFERENCE        A47537
   #authors      Boise, L.H.; Gonzalez-Garcia, M.; Postema, C.E.; Ding, L.;
                   Lindsten, T.; Turka, L.A.; Mao, X.; Nunez, G.; Thompson,
                   C.B.
   #journal      Cell (1993) 74:597-608
   #title        bcl-x, a bcl-2-related gene that functions as a dominant
                   regulator of apoptotic cell death.
   #accession    A47537
      ##status        preliminary
      ##molecule_type DNA
      ##residues      1-190 ##label BOI
      ##cross-references GB:L20120
SUMMARY          #length 190  #molecular-weight 21467  #checksum 5509

  Query Match           17.4%;  Score 271;  DB 13;  Length 190;
  Best Local Similarity  28.1%;  Pred. No. 1.24e-23;
  Matches    34;  Conservative   36;  Mismatches  49;  Indels   2;  Gaps   2;

Db      70 slevheivrasdvrqalrdagdefelryrrafsdltsqlhitpgtayqsfeqvvnelfhd 129
           :| ::     ::| : |   ||::: ||   |  :  :|: |:  ||: |  : : ||
Qy      62 TLPLQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENAYEYFTKIASSLFES 121

Db     130 gvnwgrivaffsfggalcvesvdkemrvlvgrivswmttyltdh-ldpwiqenggwvrta 188
           |:||||:||:::|:  | :   :: :  ::|::   :  ::  | :  || : |||| :|
Qy     122 GINWGRVVALLGFSYRLALHIYQRGLTGFLGQVTRFVVDFMLHHCIARWIAQRGGWV-AA 180

Db     189 l 189
           |
Qy     181 L 181


RESULT    4
ENTRY            B37332     #type complete
TITLE            transforming protein (bcl-2-beta) - chicken
ORGANISM         #formal_name Gallus gallus #common_name chicken
DATE             03-Mar-1993 #sequence_revision 03-Mar-1993 #text_change
                   15-Jun-1996
ACCESSIONS       B37332
REFERENCE        A37332
   #authors      Eguchi, Y.; Ewert, D.L.; Tsujimoto, Y.
   #journal      Nucleic Acids Res. (1992) 20:4187-4192
   #title        Isolation and characterization of the chicken bcl-2 gene:
                   expression in a variety of tissues including lymphoid and
                   neuronal organs in adult and embryo.
   #accession    B37332
      ##status        preliminary; nucleic acid sequence not shown; not
                        compared with conceptual translation
      ##molecule_type DNA
      ##residues      1-216 ##label EGU
      ##cross-references EMBL:D11382
CLASSIFICATION   #superfamily bcl transforming protein
SUMMARY          #length 216  #molecular-weight 23492  #checksum 3596

  Query Match           17.2%;  Score 267;  DB 6;  Length 216;
  Best Local Similarity  28.5%;  Pred. No. 5.03e-23;
  Matches    39;  Conservative   39;  Mismatches  55;  Indels   4;  Gaps   4;

Db      60 hhrpeppgsaaasevppae-glrpapp-g-vhlalrqagdefsrryqrdfaqmsgqlhlt 116
           ::  |: |:||:::       |:|::  | |   |   ||:::|||: :|  |  :|: |
Qy      44 QQEQEAEGAAAPADPEMVTLPLQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPT 103

Db     117 pftahgrfvavveelfrdgvnwgrivaffefggvmcvesvnremsplvdniatwmteyln 176
           :  |   |  :   ||  |:||||:||:: |:  : :    | :: ::  ::  : :::
Qy     104 AENAYEYFTKIASSLFESGINWGRVVALLGFSYRLALHIYQRGLTGFLGQVTRFVVDFML 163

Db     177 rh-lhnwiqdnggwvra 192
           :| :  || : |||| |
Qy     164 HHCIARWIAQRGGWVAA 180


RESULT    5
ENTRY            S51761     #type complete
TITLE            BCL-X protein - rat
ORGANISM         #formal_name Rattus norvegicus #common_name Norway rat
```

```
DATE            07-May-1995 #sequence_revision 01-Sep-1995 #text_change
                  01-Dec-1995
ACCESSIONS      S51761; S51762
REFERENCE       S51761
   #authors     Michaelidis, T.M.
   #submission  submitted to the EMBL Data Library, November 1994
   #accession   S51761
      ##status        preliminary
      ##molecule_type DNA
      ##residues      1-233 ##label MIC
      ##cross-references EMBL:X82537
REFERENCE       S51761
   #authors     Michaelidis, T.M.
   #submission  submitted to the EMBL Data Library, November 1994
   #accession   S51762
      ##status        preliminary
      ##molecule_type DNA
      ##residues      1-125,189-233 ##label MI2
      ##cross-references EMBL:X82537
GENETICS
   #introns     125/3
SUMMARY         #length 233  #molecular-weight 26130  #checksum 6378

  Query Match           17.2%;  Score 268;  DB 14;  Length 233;
  Best Local Similarity  28.0%;  Pred. No. 3.55e-23;
  Matches    33;  Conservative   37;  Mismatches  46;  Indels   2;  Gaps   2;

Db      83 maavkqalreagdefelryrrafsdltsqlhitpgtayqsfeqvvnelfrdgvnwgriva 142
           |: | : |   ||::: ||   |  :  :|: |:  ||: |  : : ||  |:||||:||
Qy      71 MGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENAYEYFTKIASSLFESGINWGRVVA 130

Db     143 ffsfggalcvesvdkemqvlvsriaswmatylndh-lepwiqenggwdtfvdlygnna 199
           :::|:  | :   :: :  :::::  :  ::  | :  || : ||| : ::| || :
Qy     131 LLGFSYRLALHIYQRGLTGFLGQVTRFVVDFMLHHCIARWIAQRGGWVAALNL-GNGP 187


RESULT    6
ENTRY           I49056     #type complete
TITLE           bcl-x long - mouse
ORGANISM        #formal_name Mus musculus #common_name house mouse
DATE            02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change
                  15-Oct-1996
ACCESSIONS      I49056; S52866
REFERENCE       I49055
   #authors     Fang, W.; Rivard, J.J.; Mueller, D.L.; Behrens, T.W.
   #journal     J. Immunol. (1994) 153:4388-4398
   #title       Cloning and molecular characterization of mouse bcl-x in B
                  and T lymphocytes.
   #cross-references MUID:95052604
   #accession   I49056
      ##status        preliminary; translated from GB/EMBL/DDBJ
      ##molecule_type mRNA
      ##residues      1-233 ##label RES
      ##cross-references EMBL:U10101; NID:g506647; CDS_PID:g506648
REFERENCE       S52866
   #authors     Kamesaki, H.; Michaud, G.Y.; Takatsu, K.; Okuma, M.
   #submission  submitted to the EMBL Data Library, November 1994
   #description IL-5 inhibits anti-IgM-induced apoptosis in an immature B
                  cell line through inductin of bcl-Xl.
   #accession   S52866
      ##status        preliminary
      ##molecule_type mRNA
      ##residues      1-233 ##label KAM
      ##cross-references EMBL:X83574
SUMMARY         #length 233  #molecular-weight 26132  #checksum 5739

  Query Match           17.2%;  Score 268;  DB 14;  Length 233;
  Best Local Similarity  28.0%;  Pred. No. 3.55e-23;
  Matches    33;  Conservative   37;  Mismatches  46;  Indels   2;  Gaps   2;

Db      83 maavkqalreagdefelryrrafsdltsqlhitpgtayqsfeqvvnelfrdgvnwgriva 142
           |: | : |   ||::: ||   |  :  :|: |:  ||: |  : : ||  |:||||:||
Qy      71 MGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENAYEYFTKIASSLFESGINWGRVVA 130

Db     143 ffsfggalcvesvdkemqvlvsriaswmatylndh-lepwiqenggwdtfvdlygnna 199
           :::|:  | :   :: :  :::::  :  ::  | :  || : ||| : ::| || :
Qy     131 LLGFSYRLALHIYQRGLTGFLGQVTRFVVDFMLHHCIARWIAQRGGWVAALNL-GNGP 187


RESULT    7
ENTRY           A37332     #type complete
TITLE           transforming protein (bcl-2-alpha) - chicken
ORGANISM        #formal_name Gallus gallus #common_name chicken
DATE            03-Mar-1993 #sequence_revision 03-Mar-1993 #text_change
                  12-Apr-1995
ACCESSIONS      A37332; S35453
REFERENCE       A37332
   #authors     Eguchi, Y.; Ewert, D.L.; Tsujimoto, Y.
   #journal     Nucleic Acids Res. (1992) 20:4187-4192
   #title       Isolation and characterization of the chicken bcl-2 gene:
                  expression in a variety of tissues including lymphoid and
                  neuronal organs in adult and embryo.
   #accession   A37332
      ##status        nucleic acid sequence not shown
      ##molecule_type DNA
      ##residues      1-233 ##label EGU
      ##cross-references EMBL:D11381
GENETICS
   #introns     189/3
CLASSIFICATION  #superfamily bcl transforming protein
KEYWORDS        transforming protein
SUMMARY         #length 233  #molecular-weight 25687  #checksum 99

  Query Match           17.1%;  Score 266;  DB 6;  Length 233;
  Best Local Similarity  28.5%;  Pred. No. 7.14e-23;
  Matches    41;  Conservative   42;  Mismatches  56;  Indels   5;  Gaps   5;

Db      60 hhrpeppgsaaasevppae-glrpapp-g-vhlalrqagdefsrryqrdfaqmsgqlhlt 116
           ::  |: |:||:::       |:|::  | |   |   ||:::|||: :|  |  :|: |
Qy      44 QQEQEAEGAAAPADPEMVTLPLQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPT 103

Db     117 pftahgrfvavveelfrdgvnwgrivaffefggvmcvesvnremsplvdniatwmteyln 176
           :  |   |  :   ||  |:||||:||:: |:  : :    | :: ::  ::  : :::
Qy     104 AENAYEYFTKIASSLFESGINWGRVVALLGFSYRLALHIYQRGLTGFLGQVTRFVVDFML 163

Db     177 rh-lhnwiqdnggwdafvelygns 199
           :| :  || : ||| | ::| ||:
Qy     164 HHCIARWIAQRGGWVAALNL-GNG 186


RESULT    8
ENTRY           B47537     #type complete
TITLE           apoptosis regulator bcl-xL - human
ALTERNATE_NAMES bcl-2-related protein
CONTAINS        apoptosis regulator bcl-xS
ORGANISM        #formal_name Homo sapiens #common_name man
DATE            16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change
                  06-Sep-1996
ACCESSIONS      B47537; C47537
REFERENCE       A47537
   #authors     Boise, L.H.; Gonzalez-Garcia, M.; Postema, C.E.; Ding, L.;
                  Lindsten, T.; Turka, L.A.; Mao, X.; Nunez, G.; Thompson,
                  C.B.
   #journal     Cell (1993) 74:597-608
   #title       bcl-x, a bcl-2-related gene that functions as a dominant
                  regulator of apoptotic cell death.
   #accession   B47537
      ##status        nucleic acid sequence not shown; translated from
                       GB/EMBL/DDBJ
      ##molecule_type mRNA
      ##residues      1-233 ##label BOI
      ##cross-references GB:L20121; CDS_PID:Q07817
   #accession   C47537
      ##status        nucleic acid sequence not shown; translated from
                       GB/EMBL/DDBJ
      ##molecule_type mRNA
```

```
     ##residues      1-69,'G',71-125,189-233 ##label BO2
     ##cross-references GB:L20122; CDS_PID:g623237
GENETICS
   #gene         GDB:BCL2L
     ##cross-references GDB:228079
KEYWORDS         alternative splicing; apoptosis
FEATURE
   1-233              #product apoptosis regulator bcl-xL #status predicted
                        #label MAT\
   1-125,189-233      #product apoptosis regulator bcl-xS #status predicted
                        #label MA2
SUMMARY          #length 233  #molecular-weight 26063  #checksum 5340

  Query Match           17.0%;  Score 264;  DB 13;  Length 233;
  Best Local Similarity  28.0%;  Pred. No. 1.43e-22;
  Matches    33;  Conservative   37;  Mismatches  46;  Indels   2;  Gaps   2;

Db     83 maavkqalreagdefelryrrafsdltsqlhitpgtayqsfeqvvnelfrdgvnwgriva 142
          |: | : |   ||::: ||   |  :  :|: |:  ||: |  : : ||  |:||||:||
Qy     71 MGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENAYEYFTKIASSLFESGINWGRVVA 130

Db    143 ffsfggalcvesvdkemqvlvsriaawmatylndh-lepwiqenggwdtfvelygnna 199
          :::|:  | :   :: :  :::::  :  ::  | :  || : ||| : ::| || :
Qy    131 LLGFSYRLALHIYQRGLTGFLGQVTRFVVDFMLHHCIARWIAQRGGWVAALNL-GNGP 187


RESULT    9
ENTRY            I49057     #type complete
TITLE            bcl-x transmembrane deleted - mouse
ORGANISM         #formal_name Mus musculus #common_name house mouse
DATE             02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change
                   02-Jul-1996
ACCESSIONS       I49057
REFERENCE        I49055
   #authors      Fang, W.; Rivard, J.J.; Mueller, D.L.; Behrens, T.W.
   #journal      J. Immunol. (1994) 153:4388-4398
   #title        Cloning and molecular characterization of mouse bcl-x in B
                   and T lymphocytes.
   #cross-references MUID:95052604
   #accession    I49057
     ##status        preliminary; translated from GB/EMBL/DDBJ
     ##molecule_type mRNA
     ##residues      1-214 ##label RES
     ##cross-references EMBL:U10102; NID:g506649; CDS_PID:g506650
GENETICS
   #note         gene name bcl-x long
SUMMARY          #length 214  #molecular-weight 23900  #checksum 9730

  Query Match           16.8%;  Score 261;  DB 14;  Length 214;
  Best Local Similarity  28.0%;  Pred. No. 4.08e-22;
  Matches    30;  Conservative   33;  Mismatches  43;  Indels   1;  Gaps   1;

Db     83 maavkqalreagdefelryrrafsdltsqlhitpgtayqsfeqvvnelfrdgvnwgriva 142
          |: | : |   ||::: ||   |  :  :|: |:  ||: |  : : ||  |:||||:||
Qy     71 MGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENAYEYFTKIASSLFESGINWGRVVA 130

Db    143 ffsfggalcvesvdkemqvlvsriaswmatylndh-lepwiqenggw 188
          :::|:  | :   :: :  ::::::  :  ::  | :  || : |||
Qy    131 LLGFSYRLALHIYQRGLTGFLGQVTRFVVDFMLHHCIARWIAQRGGW 177


RESULT   10
ENTRY            TVHUB1     #type complete
TITLE            transforming protein bcl-2-beta - human
ORGANISM         #formal_name Homo sapiens #common_name man
DATE             31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change
                   06-Sep-1996
ACCESSIONS       B29409
REFERENCE        A29409
   #authors      Tsujimoto, Y.; Croce, C.M.
   #journal      Proc. Natl. Acad. Sci. U.S.A. (1986) 83:5214-5218
   #title        Analysis of the structure, transcripts, and protein products
                   of bcl-2, the gene involved in human follicular lymphoma.
   #cross-references MUID:86259760
   #accession    B29409
     ##molecule_type mRNA
     ##residues      1-205 ##label TSU
GENETICS
   #gene         GDB:BCL2
     ##cross-references GDB:119031
   #map_position 18q21.33-18q21.33
CLASSIFICATION   #superfamily bcl transforming protein
KEYWORDS         alternative splicing; B-cell lymphoma; follicular lymphoma;
                   transforming protein
SUMMARY          #length 205  #molecular-weight 22182  #checksum 1183

  Query Match           16.7%;  Score 260;  DB 2;  Length 205;
  Best Local Similarity  29.7%;  Pred. No. 5.77e-22;
  Matches    38;  Conservative   37;  Mismatches  49;  Indels   4;  Gaps   3;

Db     74 tpaapgaaagpalspvppvvhla--lrqagddfsrryrgdfaemssqlhltpftargrfa 131
          :|| |   : | | | : : :::  |   |||::||| ::|  |  :|: |:  |   |:
Qy     54 APADPEMVTLP-LQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENAYEYFT 112

Db    132 tvveelfrdgvnwgrivaffefggvmcvesvnremsplvdnialwmteylnrh-lhtwiq 190
          :   ||  |:||||:||:: |:  : :    | :: ::  ::  : ::: :| :  ||
Qy    113 KIASSLFESGINWGRVVALLGFSYRLALHIYQRGLTGFLGQVTRFVVDFMLHHCIARWIA 172

Db    191 dnggwvga 198
          : ||||:|
Qy    173 QRGGWVAA 180


RESULT   11
ENTRY            D37332     #type complete
TITLE            transforming protein (bcl-2-beta) - human
ORGANISM         #formal_name Homo sapiens #common_name man
DATE             03-Mar-1993 #sequence_revision 03-Mar-1993 #text_change
                   15-Jun-1996
ACCESSIONS       D37332
REFERENCE        A37332
   #authors      Eguchi, Y.; Ewert, D.L.; Tsujimoto, Y.
   #journal      Nucleic Acids Res. (1992) 20:4187-4192
   #title        Isolation and characterization of the chicken bcl-2 gene:
                   expression in a variety of tissues including lymphoid and
                   neuronal organs in adult and embryo.
   #accession    D37332
     ##status        preliminary; nucleic acid sequence not shown; not
                       compared with conceptual translation
     ##molecule_type DNA
     ##residues      1-206 ##label EGU
CLASSIFICATION   #superfamily bcl transforming protein
SUMMARY          #length 206  #molecular-weight 22440  #checksum 5581

  Query Match           16.5%;  Score 256;  DB 6;  Length 206;
  Best Local Similarity  29.7%;  Pred. No. 2.31e-21;
  Matches    38;  Conservative   35;  Mismatches  51;  Indels   4;  Gaps   3;

Db     75 tpaapgaaagpalspvppvvhlt--lrqagddfsrryrrdfaemssqlhltpftargrfa 132
          :|| |   : | | | : : ::   |   |||::|||  :|  |  :|: |:  |   |:
Qy     54 APADPEMVTLP-LQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENAYEYFT 112

Db    133 tvveelfrdgvnwgrivaffefggvmcvesvnremsplvdnialwmteylnrh-lhtwiq 191
          :   ||  |:||||:||:: |:  : :    | :: ::  ::  : ::: :| :  ||
Qy    113 KIASSLFESGINWGRVVALLGFSYRLALHIYQRGLTGFLGQVTRFVVDFMLHHCIARWIA 172

Db    192 dnggwvga 199
          : ||||:|
Qy    173 QRGGWVAA 180


RESULT   12
ENTRY            TVHUA1     #type complete
TITLE            transforming protein bcl-2, splice form alpha - human
ORGANISM         #formal_name Homo sapiens #common_name man
DATE             31-Dec-1988 #sequence_revision 07-Jun-1996 #text_change
```

```
                 06-Sep-1996
ACCESSIONS       C37332; A29409; S02452; A24428; A27622; B27622
REFERENCE        A37332
   #authors      Eguchi, Y.; Ewert, D.L.; Tsujimoto, Y.
   #journal      Nucleic Acids Res. (1992) 20:4187-4192
   #title        Isolation and characterization of the chicken bcl-2 gene:
                   expression in a variety of tissues including lymphoid and
                   neuronal organs in adult and embryo.
   #accession    C37332
      ##status        nucleic acid sequence not shown; not compared with
                        conceptual translation
      ##molecule_type DNA
      ##residues      1-239 ##label EGU
      ##note          this report is a correction
REFERENCE        A29409
   #authors      Tsujimoto, Y.; Croce, C.M.
   #journal      Proc. Natl. Acad. Sci. U.S.A. (1986) 83:5214-5218
   #title        Analysis of the structure, transcripts, and protein products
                   of bcl-2, the gene involved in human follicular lymphoma.
   #cross-references MUID:86259760
   #accession    A29409
      ##molecule_type mRNA
      ##residues      1-95,'A',97-109,'G',111-236,'S',238-239 ##label TSU
      ##note          this sequence has been corrected in reference A37332
REFERENCE        S02452
   #authors      Seto, M.; Jaeger, U.; Hockett, R.D.; Graninger, W.; Bennett,
                   S.; Goldman, P.; Korsmeyer, S.J.
   #journal      EMBO J. (1988) 7:123-131
   #title        Alternative promoters and exons, somatic mutation and
                   deregulation of the Bcl-2--Ig fusion gene in lymphoma.
   #cross-references MUID:88196071
   #accession    S02452
      ##molecule_type mRNA
      ##residues      1-239 ##label SET
REFERENCE        A24428
   #authors      Cleary, M.L.; Smith, S.D.; Sklar, J.
   #journal      Cell (1986) 47:19-28
   #title        Cloning and structural analysis of cDNAs for bcl-2 and a
                   hybrid bcl-2/immunoglobulin transcript resulting from the t
                   (14;18) translocation.
   #cross-references MUID:87002488
   #accession    A24428
      ##molecule_type mRNA
      ##residues      1-58,'T',60-116,'R',118-239 ##label CLE
REFERENCE        A27622
   #authors      Hua, C.; Zorn, S.; Jensen, J.P.; Coupland, R.W.; Ko, H.S.;
                   Wright, J.J.; Bakhshi, A.
   #journal      Oncogene Res. (1988) 2:263-275
   #title        Consequences of the t(14;18) chromosomal translocation in
                   follicular lymphoma: deregulated expression of a chimeric
                   and mutated BCL-2 gene.
   #cross-references MUID:88217344
   #accession    A27622
      ##molecule_type mRNA
      ##residues      1-58,'T',60-239 ##label HUA
   #accession    B27622
      ##molecule_type DNA
      ##residues      1-6,'S',8-58,'T',60-128,'C',130-239 ##label HUA2
      ##note          the sequence was determined from the germline gene
COMMENT     Constitutive expression of BCL2 following t(14:18) chromosomal
              translocation is typically found in follicular lymphoma.
GENETICS
   #gene         GDB:BCL2
      ##cross-references GDB:119031
   #map_position 18q21.33-18q21.33
FUNCTION
   #description  blocks apoptosis in hematopoietic cells
CLASSIFICATION   #superfamily bcl transforming protein
KEYWORDS         alternative splicing; B-cell lymphoma; follicular lymphoma;
                   proto-oncogene; transforming protein
SUMMARY          #length 239  #molecular-weight 26266  #checksum 8323

  Query Match           16.5%;  Score 256;  DB 2;  Length 239;
  Best Local Similarity  29.3%;  Pred. No. 2.31e-21;
  Matches    41;  Conservative   37;  Mismatches  57;  Indels   5;  Gaps   4;

Db     74 tpaapgaaagpalspvppvvhlt--lrqagddfsrryrrdfaemssqlhltpftargrfa 131
          :|| |   : | | | : : ::   |   |||::||| :|  |  :|: |:  |   |:
Qy     54 APADPEMVTLP-LQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENAYEYFT 112

Db    132 tvveelfrdgvnwgrivaffefggvmcvesvnremsplvdnialwmteylnrh-lhtwiq 190
           :   ||  |:||||:||:: |:  : :    | :: ::  ::  : ::: :| :  ||
Qy    113 KIASSLFESGINWGRVVALLGFSYRLALHIYQRGLTGFLGQVTRFVVDFMLHHCIARWIA 172

Db    191 dnggwdafvel-ygpsmrpl 209
          : ||| | ::|  || :  |
Qy    173 QRGGWVAALNLGNGPILNVL 192


RESULT   13
ENTRY            TVMSB1     #type complete
TITLE            transforming protein bcl-2-beta - mouse
ORGANISM         #formal_name Mus musculus #common_name house mouse
DATE             31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change
                   02-Jun-1994
ACCESSIONS       B25960
REFERENCE        A90893
   #authors      Negrini, M.; Silini, E.; Kozak, C.; Tsujimoto, Y.; Croce,
                   C.M.
   #journal      Cell (1987) 49:455-463
   #title        Molecular analysis of mbcl-2: structure and expression of the
                   murine gene homologous to the human gene involved in
                   follicular lymphoma.
   #cross-references MUID:87187643
   #accession    B25960
      ##molecule_type DNA
      ##residues      1-199 ##label NEG
GENETICS
   #gene         BCL2
CLASSIFICATION   #superfamily bcl transforming protein
KEYWORDS         alternative splicing; transforming protein
SUMMARY          #length 199  #molecular-weight 22299  #checksum 7397

  Query Match           16.2%;  Score 251;  DB 2;  Length 199;
  Best Local Similarity  31.1%;  Pred. No. 1.30e-20;
  Matches    32;  Conservative   29;  Mismatches  41;  Indels   1;  Gaps   1;

Db     94 lrragddfsrryrrdfaemssqlhltpftargrfatvveelfrdgvnwgrivaffefggv 153
          |   |||::||| :|  |  :|: |:  |   |: :   ||  |:||||:||:: |:
Qy     78 LAIIGDDINRRYDSEFQTMLQHLQPTAENAYEYFTKIASSLFESGINWGRVVALLGFSYR 137

Db    154 mcvesvnremsplvdnialwmteylnrh-lhtwiqdnggwvga 195
          : :    | :: ::  ::  : ::: :| :  || : ||||:|
Qy    138 LALHIYQRGLTGFLGQVTRFVVDFMLHHCIARWIAQRGGWVAA 180


RESULT   14
ENTRY            I67431     #type complete
TITLE            BCL-X-Long - rat
ORGANISM         #formal_name Rattus norvegicus #common_name Norway rat
DATE             26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change
                   26-Jul-1996
ACCESSIONS       I67431
REFERENCE        I53295
   #authors      Tilly, J.L.; Tilly, K.I.; Kenton, M.L.; Johnson, A.L.
   #journal      Endocrinology (1995) 136:232-241
   #title        Expression of members of the bcl-2 gene family in the
                   immature rat ovary: equine chorionic gonadotropin-mediated
                   inhibition of granulosa cell apoptosis is associated with
                   decreased bax and constitutive bcl-2 and bcl-xlong
                   messenger ribonucleic acid levels.
   #cross-references MUID:95129487
   #accession    I67431
      ##status        preliminary; translated from GB/EMBL/DDBJ
      ##molecule_type mRNA
      ##residues      1-233 ##label RES
```

```
     ##cross-references EMBL:U34963; NID:g1004376; CDS_PID:g1004377
SUMMARY          #length 233  #molecular-weight 26122  #checksum 8310

  Query Match           16.2%;  Score 251;  DB 14;  Length 233;
  Best Local Similarity  27.6%;  Pred. No. 1.30e-20;
  Matches    32;  Conservative   34;  Mismatches  48;  Indels   2;  Gaps   2;

Db     83 maavkqalreagdefelryrrafsdltsqlhitpgtvyqsfeqvvnelfrdgvnwgriva 142
          |: | : |   ||::: ||   |  :  :|: |:   |: |  : : ||  |:||||:||
Qy     71 MGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENAYEYFTKIASSLFESGINWGRVVA 130

Db    143 sssfggalcvesvdkemqvlvsriaswmatylndh-lepwiqenggwdtfvdlygn 197
            :|:  | :   :: :  ::::::  :  ::  | :  || : ||| : ::| ||
Qy    131 LLGFSYRLALHIYQRGLTGFLGQVTRFVVDFMLHHCIARWIAQRGGWVAALNL-GN 185


RESULT   15
ENTRY            TVMSA1     #type complete
TITLE            transforming protein bcl-2-alpha - mouse
ORGANISM         #formal_name Mus musculus #common_name house mouse
DATE             31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change
                   02-Jun-1994
ACCESSIONS       A25960
REFERENCE        A90893
   #authors      Negrini, M.; Silini, E.; Kozak, C.; Tsujimoto, Y.; Croce,
                   C.M.
   #journal      Cell (1987) 49:455-463
   #title        Molecular analysis of mbcl-2: structure and expression of the
                   murine gene homologous to the human gene involved in
                   follicular lymphoma.
   #cross-references MUID:87187643
   #accession    A25960
      ##molecule_type DNA
      ##residues      1-236 ##label NEG
GENETICS
   #gene         BCL2
   #introns      192/3
CLASSIFICATION   #superfamily bcl transforming protein
KEYWORDS         alternative splicing; transforming protein
SUMMARY          #length 236  #molecular-weight 26524  #checksum 6709

  Query Match           16.2%;  Score 251;  DB 2;  Length 236;
  Best Local Similarity  30.4%;  Pred. No. 1.30e-20;
  Matches    35;  Conservative   31;  Mismatches  47;  Indels   2;  Gaps   2;

Db     94 lrragddfsrryrrdfaemssqlhltpftargrfatvveelfrdgvnwgrivaffefggv 153
          |   |||::|||  :|  |  :|: |:  |   |: :   ||  |:||||:||:: |:
Qy     78 LAIIGDDINRRYDSEFQTMLQHLQPTAENAYEYFTKIASSLFESGINWGRVVALLGFSYR 137

Db    154 mcvesvnremsplvdnialwmteylnrh-lhtwiqdnggwdafvel-ygpsmrpl 206
          : :    | :: ::  ::  : ::: :| :  || : ||| | ::|  || :  |
Qy    138 LALHIYQRGLTGFLGQVTRFVVDFMLHHCIARWIAQRGGWVAALNLGNGPILNVL 192


Search completed: Wed Aug 20 11:05:58 1997
Job time : 25 secs.
```

Release 2.1D John F. Collins, Biocomputing Research Unit.

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Wed Aug 20 11:08:18 1997; MasPar time 3.38 Seconds
298.086 Million cell updates/sec

Tabular output not generated.

Title: >US-08-320-157-21
Description: (1-88) from US08320157.pep
Perfect Score: 643
Sequence: 1 MASGQGPGPPRQECGKPALP.....APWGRWDGSSPSPGRHQPAL 88

Scoring table: PAM 150
Gap 11

Searched: 96640 seqs, 11439865 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: a-geneseq27
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20

Statistics: Mean 27.717; Variance 105.074; scale 0.264

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description	Pred. No.
1	643	100.0	88	13	R77878	Human Cdn-3.	1.98e-58
2	213	33.1	211	20	W03669	Bak-2 protein.	3.66e-12
3	213	33.1	211	13	R77876	Human Cdn-1.	3.66e-12
4	213	33.1	211	20	W03668	Bak protein.	3.66e-12
5	213	33.1	211	17	R81451	Bcl-Y apoptosis-relat	3.66e-12
6	213	33.1	211	13	R77877	Human Cdn-2.	3.66e-12
7	84	13.1	647	5	R03928	Xenopus laevis HSP (x	7.53e+00
8	83	12.9	641	8	R43004	Mature mouse sperm 70	9.13e+00
9	82	12.8	634	5	R03930	Gallus gallus HSP (ch	1.11e+01
10	81	12.6	341	18	R95830	Human interleukin-I-c	1.34e+01
11	81	12.6	459	13	R66729	Aromatic dihydrodiol/	1.34e+01
12	81	12.6	2510	14	R71007	Human neuronal calciu	1.34e+01
13	80	12.4	233	13	R64327	HSV L/ST ORF1.	1.62e+01
14	80	12.4	982	3	R13320	Murine Natural Killer	1.62e+01
15	80	12.4	1023	3	R13319	Partial Human Natural	1.62e+01
16	79	12.3	633	8	R43002	Mouse SLIP1 homologue	1.95e+01
17	78	12.1	429	8	R41746	MN protein.	2.36e+01
18	78	12.1	646	5	R03927	Rat HSP (rathsp70).	2.36e+01
19	76	11.8	501	12	R70996	Human alpha-1A adrene	3.43e+01
20	76	11.8	501	18	R90040	Human alpha-1A adrene	3.43e+01
21	76	11.8	510	7	R34674	Insulinoma-associated	3.43e+01
22	76	11.8	3398	8	R44430	eryA region polypepti	3.43e+01
23	74	11.5	487	15	R88619	Phytase protein.	4.96e+01
24	74	11.5	1540	18	R81483	BRCA1 mutant from pat	4.96e+01
25	74	11.5	1852	19	R99448	BRCA1 allele #77 tran	4.96e+01
26	74	11.5	1863	19	R99440	BRCA1 allele #8403 tr	4.96e+01
27	74	11.5	1863	15	R76641	BRCA1 protein.	4.96e+01
28	74	11.5	1863	18	R81538	BRCA1 mutant from PM1	4.96e+01
29	74	11.5	1863	18	R81542	BRCA1 mutant from PM2	4.96e+01
30	74	11.5	1863	19	R97128	BRCA1, breast and ova	4.96e+01
31	74	11.5	1863	18	R81545	BRCA1 mutant from PM2	4.96e+01
32	74	11.5	1863	18	R81481	BRCA1.	4.96e+01
33	74	11.5	1863	18	R81533	BRCA1 mutant from PM0	4.96e+01
34	74	11.5	1863	18	R81534	BRCA1 mutant from PM0	4.96e+01
35	74	11.5	1863	18	R81536	BRCA1 mutant from PM0	4.96e+01
36	74	11.5	1863	18	R81535	BRCA1 mutant from PM0	4.96e+01
37	74	11.5	1863	18	R81511	BRCA1 mutant from sam	4.96e+01
38	74	11.5	1863	18	R81514	BRCA1 mutant from sam	4.96e+01
39	74	11.5	1863	18	R81509	BRCA1 mutant from sam	4.96e+01
40	74	11.5	1863	18	R81519	BRCA1 mutant from sam	4.96e+01
41	74	11.5	1863	18	R81497	BRCA1 mutant from fam	4.96e+01
42	74	11.5	1863	18	R81485	BRCA1 mutant from pat	4.96e+01
43	74	11.5	1863	18	R81488	BRCA1 mutant from pat	4.96e+01
44	74	11.5	1863	18	R81490	BRCA1 mutant from pat	4.96e+01
45	74	11.5	1863	18	R81522	BRCA1 mutant from sam	4.96e+01

ALIGNMENTS

RESULT 1
ID R77878 standard; Protein; 88 AA.
AC R77878;
DT 21-NOV-1995 (first entry)
DE Human Cdn-3.
KW Cdn-3; apoptosis modulator; adoptive immunotherapy; therapy; HIV;
KW autoimmune disease; reperfusion injury; hepatitis, osteoporosis;
KW shock; lymphoma; eczema.
OS Homo sapiens.
PN WO9515084-A.
PD 08-JUN-1995.
PF 30-NOV-1994; U13930.
PR 30-NOV-1993; US-160067.
PR 07-OCT-1994; US-320157.
PA (LXRB-) LXR BIOTECHNOLOGY INC.
PI Barr PJ, Kiefer MC;
DR WPI; 95-215106/28.
DR N-PSDB; Q95494.
PT New nucleic acid sequences encoding Cdn apoptosis modulators - and
PT related vectors, transformed cells, proteins and antibodies, useful
PT or diagnosis and treatment e.g. of HIV infection, reperfusion injury
PT etc.
PS Disclosure; Fig.7B-C; 66pp; English.
CC Southern blot analysis of human genomic DNA and a panel of
CC human/rodent somatic cell DNAs revealed at least 3 Cdn-related
CC genes residing on chromosomes 6, 11 and 20. Cdn-3 (R77878)
CC did not contain the structural features of Cdn-1 (R77876), Cdn-2
CC (R77877) or other Bcl-2 family members.
SQ Sequence 88 AA;

Query Match 100.0%; Score 643; DB 13; Length 88;
Best Local Similarity 100.0%; Pred. No. 1.98e-58;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Db      1 masgqgpgpprqecgkpalpsaseeqvaqdmegfsaatffttisrnrrlkgrpplptqrw 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      1 MASGQGPGPPRQECGKPALPSASEEQVAQDMEGFSAATFFTTISRNRRLKGRPPLPTQRW 60

Db     61 spcpsnlaapwgrwdgsspspgrhqpal 88
          |||||||||||||||||||||||||||||
Qy     61 SPCPSNLAAPWGRWDGSSPSPGRHQPAL 88
```

RESULT 2

```
ID   W03669 standard; Protein; 211 AA.
AC   W03669;
DT   22-FEB-1997  (first entry)
DE   Bak-2 protein.
KW   Human; Bak-2; apoptosis; latency; virus replication;
KW   Epstein-Barr virus; BHRF1; fusion protein; epitope tag;
KW   drug screening; co-precipitation; ELISA; immunoassay; antibody;
KW   protein interactive trapping; virucide; antitumour; diagnostic.
OS   Homo sapiens.
PN   WO9633416-A1.
PD   24-OCT-1996.
PF   19-APR-1996; U05639.
PR   20-APR-1995; US-426529.
PA   (LXRB-) LXR BIOTECHNOLOGY INC.
PI   Barr PJ,  Kiefer MC;
DR   WPI; 96-485886/48.
DR   N-PSDB; T42139.
PT   Screening for anti-viral agents - by detecting the ability of an
PT   agent to disrupt the interaction of a Bak protein and a viral
PT   protein
PS   Disclosure; Fig 2; 24pp; English.
CC   This Bak-2 protein sequence represents a bcl-1 homologue which
CC   interacts with Epstein-Barr virus (EBV) early lytic cycle BHRF1
CC   protein, and is capable of modulating apoptosis.  The protein may
CC   be used in complete or partial form, or as an epitope tag fusion
CC   protein, in a new virucide drug screening method, which involves
CC   combination of Bak-2 protein and a viral protein (e.g. EBV BHRF1),
CC   exposure to a test compound, and monitoring for disruption of the
CC   interaction, e.g. by co-precipitation, protein interactive trapping
CC   or ELISA.  Interaction of Bak-2 and viral proteins allows viral
CC   replication or latency in the absence of apoptosis.  Compounds which
CC   inhibit the interaction may be used as virucide, antitumour or
CC   diagnostic agents.
SQ   Sequence   211 AA;

  Query Match           33.1%;  Score 213;  DB 20;  Length 211;
  Best Local Similarity  78.0%;  Pred. No. 3.66e-12;
  Matches    32;  Conservative    2;  Mismatches   6;  Indels   1;  Gaps   1;

Db      1 masgqgpgpprqecgepalpsaseeqvaqdteevfrsyvfy 41
          |||||||||||||||| |||||||||||||| |  | :  |:
Qy      1 MASGQGPGPPRQECGKPALPSASEEQVAQDMEG-FSAATFF 40


RESULT    3
ID   R77876 standard; Protein; 211 AA.
AC   R77876;
DT   21-NOV-1995  (first entry)
DE   Human Cdn-1.
KW   Cdn-1; apoptosis modulator; adoptive immunotherapy; therapy; HIV;
KW   autoimmune disease; reperfusion injury; hepatitis, osteoporosis;
KW   shock; lymphoma; eczema.
OS   Homo sapiens.
PN   WO9515084-A.
PD   08-JUN-1995.
PF   30-NOV-1994; U13930.
PR   30-NOV-1993; US-160067.
PR   07-OCT-1994; US-320157.
PA   (LXRB-) LXR BIOTECHNOLOGY INC.
PI   Barr PJ,  Kiefer MC;
DR   WPI; 95-215106/28.
DR   N-PSDB; Q95492.
PT   New nucleic acid sequences encoding Cdn apoptosis modulators  - and
PT   related vectors, transformed cells, proteins and antibodies, useful
PT   or diagnosis and treatment e.g. of HIV infection, reperfusion injury
PT   etc.
PS   Disclosure; Fig.3A-B; 66pp; English.
CC   Cdn-1 cDNA was isolated from a human heart cDNA library using a
CC   previously isolated clone as probe.  Recombinant Cdn-1 was produced
CC   in Sf9 and human colon adenocarcinoma HT29 cells.  Expression of
CC   Cdn-1 in WI-L2 lymphoblastoid cells resulted in increased cell
CC   survival in response to anti-Fas-mediated apoptosis.
SQ   Sequence   211 AA;

  Query Match           33.1%;  Score 213;  DB 13;  Length 211;
  Best Local Similarity  78.0%;  Pred. No. 3.66e-12;
  Matches    32;  Conservative    2;  Mismatches   6;  Indels   1;  Gaps   1;

Db      1 masgqgpgpprqecgepalpsaseeqvaqdteevfrsyvfy 41
          |||||||||||||||| |||||||||||||| |  | :  |:
Qy      1 MASGQGPGPPRQECGKPALPSASEEQVAQDMEG-FSAATFF 40


RESULT    4
ID   W03668 standard; Protein; 211 AA.
AC   W03668;
DT   22-FEB-1997  (first entry)
DE   Bak protein.
KW   Human; Bak; apoptosis; latency; virus replication;
KW   Epstein-Barr virus; BHRF1; fusion protein; epitope tag;
KW   drug screening; co-precipitation; ELISA; immunoassay; antibody;
KW   protein interactive trapping; virucide; antitumour; diagnostic.
OS   Homo sapiens.
PN   WO9633416-A1.
PD   24-OCT-1996.
PF   19-APR-1996; U05639.
PR   20-APR-1995; US-426529.
PA   (LXRB-) LXR BIOTECHNOLOGY INC.
PI   Barr PJ,  Kiefer MC;
DR   WPI; 96-485886/48.
DR   N-PSDB; T42138.
PT   Screening for anti-viral agents - by detecting the ability of an
PT   agent to disrupt the interaction of a Bak protein and a viral
PT   protein
PS   Disclosure; Fig 1; 24pp; English.
CC   This Bak protein sequence represents a bcl-1 homologue which
CC   interacts with Epstein-Barr virus (EBV) early lytic cycle BHRF1
CC   protein, and is capable of modulating apoptosis.  The protein may
CC   be used in complete or partial form, or as an epitope tag fusion
CC   protein, in a new virucide drug screening method, which involves
CC   combination of Bak protein and a viral protein (e.g. EBV BHRF1),
CC   exposure to a test compound, and monitoring for disruption of the
CC   interaction, e.g. by co-precipitation, protein interactive trapping
CC   or ELISA.  Interaction of Bak and viral proteins allows viral
CC   replication or latency in the absence of apoptosis.  Compounds which
CC   inhibit the interaction may be used as virucide, antitumour or
CC   diagnostic agents.
SQ   Sequence   211 AA;

  Query Match           33.1%;  Score 213;  DB 20;  Length 211;
  Best Local Similarity  78.0%;  Pred. No. 3.66e-12;
  Matches    32;  Conservative    2;  Mismatches   6;  Indels   1;  Gaps   1;

Db      1 masgqgpgpprqecgepalpsaseeqvaqdteevfrsyvfy 41
          |||||||||||||||| |||||||||||||| |  | :  |:
Qy      1 MASGQGPGPPRQECGKPALPSASEEQVAQDMEG-FSAATFF 40


RESULT    5
ID   R81451 standard; Protein; 211 AA.
AC   R81451;
DT   02-JUL-1996  (first entry)
DE   Bcl-Y apoptosis-related protein.
KW   Bcl-Y; apoptosis; cell proliferation; cell death; diagnosis;
KW   therapy.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   Domain          188..205
FT   /label= C-terminal_domain
FT   /note= "putative membrane localisation sequence"
PN   WO9605232-A1.
PD   22-FEB-1996.
PF   09-AUG-1995; U10103.
PR   09-AUG-1994; US-287427.
PR   11-OCT-1994; US-321071.
PA   (IMMU-) IMMUNOGEN INC.
```

```
PI   Chittenden TD;
DR   WPI; 96-139648/14.
DR   N-PSDB; T17375.
PT   New isolated human Bcl-Y protein - used to develop prods. for
PT   treating disorders characterised by inappropriate cell proliferation
PT   or cell death
PS   Claim 3; Fig 4; 100pp; English.
CC   Bcl-Y protein (R81451) is a member of the Bcl-2 family and can
CC   induce apoptosis in cells and function as a negative regulator of
CC   Bcl-2 function.  Bcl-Y mRNA was detected in all human tumour cell
CC   lines examined and is also widely expressed in primary human
CC   tissues.  It can be obtd. by expression of a full-length cDNA
CC   clone (T17375) in pref. mammalian host cells.  Bcl-Y can be used to
CC   develop prods. for treating disorders associated with inappropriate
CC   cell proliferation or cell death, and to raise antibodies used for
CC   the diagnosis or monitoring of such disorders.
SQ   Sequence   211 AA;

  Query Match           33.1%;  Score 213;  DB 17;  Length 211;
  Best Local Similarity  78.0%;  Pred. No. 3.66e-12;
  Matches    32;  Conservative    2;  Mismatches   6;  Indels   1;  Gaps   1;

Db       1 masgqgpgpprqecgepalpsaseeqvaqdteevfrsyvfy 41
           |||||||||||||||| |||||||||||||| |  | :  |:
Qy       1 MASGQGPGPPRQECGKPALPSASEEQVAQDMEG-FSAATFF 40


RESULT    6
ID   R77877 standard; Protein; 211 AA.
AC   R77877;
DT   21-NOV-1995  (first entry)
DE   Human Cdn-2.
KW   Cdn-2; apoptosis modulator; adoptive immunotherapy; therapy; HIV;
KW   autoimmune disease; reperfusion injury; hepatitis, osteoporosis;
KW   shock; lymphoma; eczema.
OS   Homo sapiens.
PN   WO9515084-A.
PD   08-JUN-1995.
PF   30-NOV-1994; U13930.
PR   30-NOV-1993; US-160067.
PR   07-OCT-1994; US-320157.
PA   (LXRB-) LXR BIOTECHNOLOGY INC.
PI   Barr PJ,  Kiefer MC;
DR   WPI; 95-215106/28.
DR   N-PSDB; Q95493.
PT   New nucleic acid sequences encoding Cdn apoptosis modulators  - and
PT   related vectors, transformed cells, proteins and antibodies, useful
PT   or diagnosis and treatment e.g. of HIV infection, reperfusion injury
PT   etc.
PS   Disclosure; Fig.5D-E; 66pp; English.
CC   Cdn-2 cDNA was isold. from a human placental genomic library
CC   using a 950 bp fragment of Cdn-1 cDNA.  Expression of Cdn-2
CC   in mouse progenitor B-cell FL5.12 cells decreased IL-3-induced
CC   apoptosis.  The Cdn-2 protein displayed 97% amino acid identity
CC   with Cdn-1 (R77876).
SQ   Sequence   211 AA;

  Query Match           33.1%;  Score 213;  DB 13;  Length 211;
  Best Local Similarity  78.0%;  Pred. No. 3.66e-12;
  Matches    32;  Conservative    2;  Mismatches   6;  Indels   1;  Gaps   1;

Db       1 masgqgpgpprqecgepalpsaseeqvaqdteevfrsyvfy 41
           |||||||||||||||| |||||||||||||| |  | :  |:
Qy       1 MASGQGPGPPRQECGKPALPSASEEQVAQDMEG-FSAATFF 40


RESULT    7
ID   R03928 standard; Protein; 647 AA.
AC   R03928;
DT   30-AUG-1990  (first entry)
DE   Xenopus laevis HSP (x170).
KW   Hsp70; heat shock protein; Trypanosoma cruzi; vaccines; HSP.
OS   Xenopus laevis.
PN   WO9002564-A.
PD   22-MAR-1990.
PF   12-SEP-1989; 03955.
PR   12-SEP-1988; US-243474.
PA   (CODO-) Codon.
PI   Dragon E, Faulds D, Sias S.
DR   WPI; 90-115820/15.
PT   Proteins homologous to heat shock proteins from Trypanosoma cruzi - used
PT   in vaccines and diagnosis for species of eg Mycoplasma or Mycobacteria.
PS   Disclosure; Fig 2.1-2.14; 86pp; English.
CC   Fig. 2 provides an alignment of heat shock proteins from a variety
CC   of organisms: 1.  M.hyopneumoniae (Mhyhsp70 - R03922);
CC   2.  Bacillus megaterium (Bmehsp70 - R03923);
CC   3.  E. coli (dnaK - R03924);
CC   4.  T. cruzi (tc70kd - R03925);
CC   5.  T. cruzi (R03926);
CC   6.  Rat rattus (rathsp70 - R03927);
CC   7.  Xenopus laevis (x170 - R03928);
CC   8.  Homo sapiens (humhsp70 - R03929);
CC   9.  Gallus gallus (chkhsp70 - R03930);
CC   10. Zea mays (mzehsp70 - R03931);
CC   11. Serratia marcescens (smahsp70 - R03932).
CC   The proteins having homology to hsp's of T. cruzi can be used in
CC   vaccines and diagnosis involving e.g. Trypanosoma, Mycoplasma and
CC   Mycobacteria species.
SQ   Sequence   647 AA;

  Query Match           13.1%;  Score 84;  DB 5;  Length 647;
  Best Local Similarity  26.3%;  Pred. No. 7.53e+00;
  Matches    10;  Conservative   12;  Mismatches  15;  Indels   1;  Gaps   1;

Db     265 rtacdrakrtlssssqasieidslfegidfytaitrar 302
           |  | ::    :|  | : :::: | :  |:|:|:| |
Qy      11 RQECGKPALPSASEEQVAQDMEG-FSAATFFTTISRNR 47


RESULT    8
ID   R43004 standard; protein; 641 AA.
AC   R43004;
DT   20-MAY-1994  (first entry)
DE   Mature mouse sperm 70kD heat shock protein.
KW   Sulphoglycolipid immobilising protein 1; sperm plasma membrane;
KW   HSC70B; mammalian; infertility; mycoplasma; HSP70.
OS   Mus musculus.
FH   Key             Location/Qualifiers
FT   Region          1..385
FT   /label= 44kD_ATPase_fragment
FT   /note= "peptides comprising an intact domain from
FT   the ATPase fragment of hsp70 are claimed;
FT   the peptides are homologous to sequences
FT   conserved between SLIP1 and 74.5kD
FT   mycoplasma protein"
FT   Domain          1..39
FT   /label= IA
FT   /note= "part"
FT   Domain          40..115
FT   /label= IB
FT   Domain          116..188
FT   /label= IA
FT   /note= "part"
FT   Domain          189..228
FT   /label= IIA
FT   /note= "part"
FT   Domain          229..306
FT   /label= IIB
FT   Domain          307..385
FT   /label= IIA
FT   /note= "part"
PN   WO9321954-A.
PD   11-NOV-1993.
PF   22-APR-1993; U03816.
PR   24-APR-1992; US-873961.
PA   (BERL-) BERLEX LAB INC.
```

```
PA   (OTTA-) OTTAWA CIVIC HOSPITAL.
PI   Faulds DH,  Lingwood CA,  Tanphaichitr N;
DR   WPI; 93-368422/46.
PT   Mammalian fertilisation decrease for detecting and treating
PT   infertility - using sulpho glyco lipid immobilising protein
PT   1-sulphated-glyco moiety interfering compsn., for mycoplasma
PT   infection treatment
PS   Claim 2 and Clalim 17; Page 60-62; 77pp; English.
CC   The likelihood of mammalian fertilisation is decreased by contacting
CC   a gamete with a sulphoglycolipid immobilising protein 1 (SLIP1)/
CC   sulphated glyco-moiety interfering composition. The interfering
CC   compsn. is e.g. the heat shock 70kD protein, SLIP1 (or analogues
CC   such as the mouse SLIP1 analogue "HSC70B" comprising the amino acid
CC   sequence R43002) or the 74.5kD mycoplasma protein (R43003).
SQ   Sequence   641 AA;

  Query Match           12.9%;  Score 83;  DB 8;  Length 641;
  Best Local Similarity  26.3%;  Pred. No. 9.13e+00;
  Matches    10;  Conservative   12;  Mismatches  15;  Indels   1;  Gaps   1;

Db    264 rtacerakrtlssstqasleidslfegidfytsitrar 301
          |  | ::    :|  | : :::: | :  |:|:|:| |
Qy     11 RQECGKPALPSASEEQVAQDMEG-FSAATFFTTISRNR 47


RESULT    9
ID   R03930 standard; Protein; 634 AA.
AC   R03930;
DT   30-AUG-1990  (first entry)
DE   Gallus gallus HSP (chkhsp70).
KW   Hsp70; heat shock protein; Trypanosoma cruzi; vaccines; HSP.
OS   Gallus gallus.
FH   Key             Location/Qualifiers
FT   Misc_difference 634
FT   /note= "residue given as "O" in specification"
PN   WO9002564-A.
PD   22-MAR-1990.
PF   12-SEP-1989; 03955.
PR   12-SEP-1988; US-243474.
PA   (CODO-) Codon.
PI   Dragon E, Faulds D, Sias S.
DR   WPI; 90-115820/15.
PT   Proteins homologous to heat shock proteins from Trypanosoma cruzi - used
PT   in vaccines and diagnosis for species of eg Mycoplasma or Mycobacteria.
PS   Disclosure; Fig 2.1-2.14; 86pp; English.
CC   According to the legend of Fig 2, the G. gallus HSP sequence has
CC   635 amino acid residues, the sequence itself has only 634,
CC   including "O" (?) at position 634.
CC   Fig. 2 provides an alignment of heat shock proteins from a variety
CC   of organisms: 1.  M.hyopneumoniae (Mhyhsp70 - R03922);
CC   2.  Bacillus megaterium (Bmehsp70 - R03923);
CC   3.  E. coli (dnaK - R03924);
CC   4.  T. cruzi (tc70kd - R03925);
CC   5.  T. cruzi (R03926);
CC   6.  Rat rattus (rathsp70 - R03927);
CC   7.  Xenopus laevis (x170 - R03928);
CC   8.  Homo sapiens (humhsp70 - R03929);
CC   9.  Gallus gallus (chkhsp70 - R03930);
CC   10. Zea mays (mzehsp70 - R03931);
CC   11. Serratia marcescens (smahsp70 - R03932).
CC   The proteins having homology to hsp's of T. cruzi can be used in
CC   vaccines and diagnosis involving e.g. Trypanosoma, Mycoplasma and
CC   Mycobacteria species.
SQ   Sequence   634 AA;

  Query Match           12.8%;  Score 82;  DB 5;  Length 634;
  Best Local Similarity  26.3%;  Pred. No. 1.11e+01;
  Matches    10;  Conservative   12;  Mismatches  15;  Indels   1;  Gaps   1;

Db    267 rtacerarrtlssstqasieidslfegidfytsitrar 304
          |  | ::    :|  | : :::: | :  |:|:|:| |
Qy     11 RQECGKPALPSASEEQVAQDMEG-FSAATFFTTISRNR 47


RESULT   10
ID   R95830 standard; Protein; 341 AA.
AC   R95830;
DT   28-OCT-1996  (first entry)
DE   Human interleukin-I-converting enzyme-like apoptosis protease-3.
KW   ICE-LAP-3; interleukin-I-converting enzyme-like apoptosis protease;
KW   enzyme; Alzheimer's disease; Parkinson's disease; septic shock;
KW   head injury; rheumatoid arthritis.
OS   Homo sapiens.
PN   WO9613603-A1.
PD   09-MAY-1996.
PF   06-JUN-1995; U07235.
PR   01-NOV-1994; US-334251.
PA   (HUMA-) HUMAN GENOME SCI INC.
PI   Hastings GA,  He W,  Hudson PL,  Rosen CA;
DR   WPI; 96-239509/24.
DR   N-PSDB; T15276.
PT   Human ICE-LAP-3 and -4 DNA and protein  - useful in the diagnosis
PT   and treatment of Alzheimer's disease, Parkinson's disease,
PT   rheumatoid arthritis, septic shock and head injury
PS   Claim 1; Page 91-92; 67pp; English.
CC   This ICE-LAP-3 protein may be used therapeutically, e.g. as an
CC   antitumor or antiviral agent and to control embryonic development
CC   and tissue homeostasis.  The protein can also be used to treat
CC   immunosuppression disorders, such as AIDS, by targeting virus
CC   infected cells for cell death.  The DNA may find use in gene
CC   therapy applications.
SQ   Sequence   341 AA;

  Query Match           12.6%;  Score 81;  DB 18;  Length 341;
  Best Local Similarity  40.0%;  Pred. No. 1.34e+01;
  Matches    12;  Conservative    8;  Mismatches   6;  Indels   4;  Gaps   3;

Db     49 lkgpqtlaaerrretvpvpaalp-pwerwq 77
          |||   |:::|:  : | |: |: || ||:
Qy     49 LKGRPPLPTQRW--S-PCPSNLAAPWGRWD 75


RESULT   11
ID   R66729 standard; Protein; 459 AA.
AC   R66729;
DT   02-OCT-1995  (first entry)
DE   Aromatic dihydrodiol/catechol deoxygenase #1.
KW   Dihydrodiol; catechol; microbial fermentation; E.coli; aromatic;
KW   deoxygenase; hydrocarbon; engineering plastic; polymerisation inhibitor;
KW   synthetic resin; drug synthesis.
OS   Pseudomonas fluorescens SD805.
PN   J06343474-A.
PD   20-DEC-1994.
PF   10-JUN-1993; 165163.
PR   10-JUN-1993; JP-165163.
PA   (SHOW ) SHOWA DENKO KK.
DR   WPI; 95-069306/10.
DR   N-PSDB; Q79907.
PT   Prepn of an aromatic cyclic di:hydro:diol cpd and a catechol cpd
PT   - using microbial cells transformed by DNA coding for enzyme
PT   protein sequences
PS   Claim 1; Page 8-9; 13pp; Japanese.
CC   The enzymes (R66729-34), encoded by genes Q79907-12, are involved in the
CC   synthesis of dihydrodiol and catechol compounds by microbial fermentation
CC   in E.coli at normal temps. and pressures from aromatic hydrocarbons.
CC   The sequence presented here is of an aromatic deoxygenase of 459 amino
CC   acids.  The aromatic dihydrodiol can be used as a raw material for
CC   engineering plastics and the catechol compound is used as a
CC   polymerisation inhibitor of synthetic resin and as an intermediate in
CC   drug synthesis.
SQ   Sequence   459 AA;

  Query Match           12.6%;  Score 81;  DB 13;  Length 459;
  Best Local Similarity  30.0%;  Pred. No. 1.34e+01;
  Matches    18;  Conservative   13;  Mismatches  26;  Indels   3;  Gaps   3;
```

```
Db    298 tkgpaaerakerlgk-vlp-a-drmvaqhmtifptcsflpgintvrtwhprgpneievws 354
          : | ::  :::  ||  || | :  ||| |  |:: :|:  |:  |    | |   : ||
Qy      2 ASGQGPGPPRQECGKPALPSASEEQVAQDMEGFSAATFFTTISRNRRLKGRPPLPTQRWS 61


RESULT   12
ID   R71007 standard; Protein; 2510 AA.
AC   R71007;
DT   01-DEC-1995  (first entry)
DE   Human neuronal calcium channel subunit alpha 1A-1.
KW   Calcium channel subunit; antagonist; agonist; diagnosis;
KW   Lambert Eaton Syndrome.
OS   Homo sapiens.
PN   WO9504822-A.
PD   16-FEB-1995.
PF   11-AUG-1994; U09230.
PR   11-AUG-1993; US-105536.
PR   05-NOV-1993; US-149097.
PA   (SALK ) SALK INST BIOTECHNOLOGY IND ASSOC.
PI   Ellis SB,  Gillespie A,  Harpold MM,  Mccue AF,  Williams ME;
DR   WPI; 95-090900/12.
DR   N-PSDB; Q84659.
PT   DNA encoding human calcium channel sub-unit(s) - used for
PT   developing prods. for studying calcium channels, e.g. for
PT   obtaining agonists and antagonists
PS   Claim 34; Page 178-190; 285pp; English.
CC   The primary transceipt of the alpha 1A subunit gene is
CC   alternatively spliced to yield at least two variant mRNAs.
CC   One form, alpha 1A-1 is given in Q84659/R71007, and the other,
CC   alpha 1A-2 is given in Q84660/R71008. Alpha 1A-2 differs from
CC   alpha 1A-1 encoding sequence at the 3' end in that it lacks a
CC   5 nt. sequence. This deletion shifts the reading frame and
CC   introduces a translation termination codon resulting in an
CC   alpha 1A-2 coding sequence that encodes a shorter alpha 1A
CC   subunit than that encoded by alpha 1A-1. DNA doncding alpha
CC   1A subunits can be isolated using all or a portion of the DNA
CC   having sequence Q84661, Q85659 or Q84660 or DNA obtd. from
CC   the phage lysate of an E. coli host contg. DNA encoding an alpha
CC   1A subunit that has been deposited in the ATCC under accession
CC   no. 75293. The DNA is such a phage includes the DNA fragment
CC   having the sequence in Q84661 which selectively hybridises under
CC   conditions of high stringency to DNA encoding alpha 1A DNA but not
CC   to DNA encoding alpha 1B.
SQ   Sequence   2510 AA;

  Query Match           12.6%;  Score 81;  DB 14;  Length 2510;
  Best Local Similarity  31.8%;  Pred. No. 1.34e+01;
  Matches    14;  Conservative   12;  Mismatches  18;  Indels   0;  Gaps   0;

Db   2235 aqerpdhgrarardqrwsrspsegrehmahrqgsssvsgspaps 2278
          :::|   ||:    ||||  ||:     ::::|||:.:|   |:
Qy     44 SRNRRLKGRPPLPTQRWSPCPSNLAAPWGRWDGSSPSPGRHQPA 87


RESULT   13
ID   R64327 standard; Protein; 233 AA.
AC   R64327;
DT   04-AUG-1995  (first entry)
DE   HSV L/ST ORF1.
KW   HSV; junction-spanning transcript; L/ST; therapy; virucide.
OS   Herpes simplex virus type 1.
PN   WO9428156-A.
PD   08-DEC-1994.
PF   20-MAY-1994; U05770.
PR   20-MAY-1993; US-065146.
PA   (DAND ) DANA FARBER CANCER INST INC.
PI   Schaffer PA,  Yeh L;
DR   WPI; 95-022825/03.
DR   N-PSDB; Q76209.
PT   Herpes Simplex Virus (HSV) specific junction spanning transcript
PT   - for inhibiting HSV L/ST synthesis, in the treatment of HSV
PT   infection.
PS   Disclosure; Page 34-35; 64pp; English.
CC   In HSV-1, 4 intron-less open reading frames (ORFs) are present
CC   within the sequence specifying the L/ST junction-spanning
CC   transcript.  ORF1-4 are given in Q76209-12, and encoded
CC   proteins in R64327-30.  Antibodies raised against such proteins
CC   may be used to identify therapeutic agents.
SQ   Sequence   233 AA;

  Query Match           12.4%;  Score 80;  DB 13;  Length 233;
  Best Local Similarity  32.5%;  Pred. No. 1.62e+01;
  Matches    13;  Conservative   12;  Mismatches  14;  Indels   1;  Gaps   1;

Db    75 arrrrrrarcaavtrrrrarrggrrkgreggwegsapppg 114
         :| || ::| :  |:|: :  :     : | |:||:|:||
Qy    44 SRNRRLKGRPPLPTQRWSPCPSNLAAPW-GRWDGSSPSPG 82


RESULT   14
ID   R13320 standard; Protein; 982 AA.
AC   R13320;
DT   22-OCT-1991  (first entry)
DE   Murine Natural Killer receptor.
KW   NK; cytotoxic drugs; tumour cell.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   Protein         8..982
FT   /label= mature_NK_receptor
FT   Peptide         1..7
FT   /label= partial_signal_peptide
PN   US7535206-A.
PD   09-JUL-1991.
PF   08-JUN-1990; 143578.
PR   08-JUN-1990; US-535206.
PA   (USSH ) NAT INST OF HEALTH.
PI   Ortaldo J, Young H, Anderson S;
DR   WPI; 91-245694/33.
DR   N-PSDB; Q13115.
PT   DNA encoding a natural killer cell receptor - used to develop
PT   prods. for the immuno-detection and immuno-therapy of tumours
PS   Disclosure; Fig 2; 30pp; English.
CC   Overlapping clones, which make up the cDNA sequence from which this
CC   sequence was deduced, were isolated from a murine PBL cDNA library
CC   prepared in lambda gt10. The purified protein can mediate the
CC   cytolytic activity of mammalian cells. It specifically distinguishes
CC   tumour cells making it a candidate for the development of products
CC   for the immunodetection and immunotherapy of tumours.
CC   See also Q13114.
SQ   Sequence   982 AA;

  Query Match           12.4%;  Score 80;  DB 3;  Length 982;
  Best Local Similarity  38.2%;  Pred. No. 1.62e+01;
  Matches    13;  Conservative    9;  Mismatches  8;  Indels   4;  Gaps   4;

Db   377 qrlrayrpp-sgekwsk-gdklsdpcssrwders 408
          :||:: ||| : ::||    :|: |  :|||  |
Qy    47 RRLKG-RPPLPTQRWSPCPSNLAAP-WGRWDGSS 78


RESULT   15
ID   R13319 standard; Protein; 1023 AA.
AC   R13319;
DT   22-OCT-1991  (first entry)
DE   Partial Human Natural Killer receptor.
KW   NK; cytotoxic drugs; tumour cell.
OS   Homo sapiens.
PN   US7535206-A.
PD   09-JUL-1991.
PF   08-JUN-1990; 143578.
PR   08-JUN-1990; US-535206.
PA   (USSH ) NAT INST OF HEALTH.
PI   Ortaldo J, Young H, Anderson S;
DR   WPI; 91-245694/33.
DR   N-PSDB; Q13114.
PT   DNA encoding a natural killer cell receptor - used to develop
```

PT prods. for the immuno-detection and immuno-therapy of tumours
PS Disclosure; Fig 1; 30pp; English.
CC Overlapping clones, which make up the cDNA sequence from which this
CC sequence was deduced, were isolated from a cDNA library prepared
CC from human NK cells purified from human peripheral blood. The
CC purified protein can mediate the cytolytic activity of mammalian
CC cells. It specifically distinguishes tumour cells making it a
CC candidate for the development of products for the immunodetection
CC and immunotherapy of tumours. See also Q13115.
SQ Sequence 1023 AA;

Query Match 12.4%; Score 80; DB 3; Length 1023;
Best Local Similarity 38.2%; Pred. No. 1.62e+01;
Matches 13; Conservative 9; Mismatches 8; Indels 4; Gaps 4;

```
Db    410 qrlrayrpp-sgekwsk-gdklsdpcssrwders 441
          :||:: ||| : ::||    :|: |  :|||  |
Qy     47 RRLKG-RPPLPTQRWSPCPSNLAAP-WGRWDGSS 78
```

Search completed: Wed Aug 20 11:08:43 1997
Job time : 25 secs.

****

MPSRCH (TM)

****

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Wed Aug 20 11:09:01 1997; MasPar time 5.05 Seconds
503.601 Million cell updates/sec
Tabular output not generated.

Title: >US-08-320-157-21
Description: (1-88) from US08320157.pep
Perfect Score: 643
Sequence: 1 MASGQGPGPPRQECGKPALP.....APWGRWDGSSPSPGRHQPAL 88

Scoring table: PAM 150
Gap 11

Searched: 91006 seqs, 28888923 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: pir51
1:ann1 2:ann2 3:ann3 4:ann4 5:unann1 6:unann2 7:unann3
8:unann4 9:unann5 10:unann6 11:unann7 12:unann8
13:unann9 14:unann10 15:unenc 16:unrev

Statistics: Mean 37.327; Variance 80.459; scale 0.464

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description	Pred. No.
1	213	33.1	211	13	S58875	cdn-2 protein - huma	9.18e-20
2	213	33.1	211	13	S58873	Bak protein - human	9.18e-20
3	95	14.8	196	9	A40623	heat shock protein Y	1.35e-01
4	89	13.8	596	13	S37577	Shb protein - human	7.82e-01
5	86	13.4	643	7	S09036	heat shock protein 7	1.84e+00
6	85	13.2	488	10	S18156	globulin 1 - eastern	2.43e+00
7	85	13.2	651	7	A36333	heat shock cognate p	2.43e+00
8	84	13.1	636	7	A48872	70k heat shock cogna	3.21e+00
9	84	13.1	639	7	S20139	heat shock protein S	3.21e+00
10	84	13.1	640	7	S37394	heat shock cognate p	3.21e+00
11	84	13.1	642	3	HHBYA1	heat shock protein S	3.21e+00
12	84	13.1	643	7	S25585	heat shock protein 7	3.21e+00
13	84	13.1	647	3	HHXL70	heat shock protein 7	3.21e+00
14	84	13.1	655	7	S18349	heat shock protein 7	3.21e+00
15	84	13.1	1878	8	S26369	genome polyprotein -	3.21e+00
16	84	13.1	2142	13	B35098	MHC class III histoc	3.21e+00
17	83	12.9	208	16	B44261	heat shock protein 7	4.23e+00
18	83	12.9	281	9	B47092	copy control protein	4.23e+00
19	83	12.9	284	12	JC1495	regulatory protein G	4.23e+00
20	83	12.9	462	5	S10235	alpha-L-fucosidase (	4.23e+00
21	83	12.9	638	7	S31766	heat shock protein 7	4.23e+00
22	83	12.9	640	7	A29160	heat shock protein,	4.23e+00
23	83	12.9	641	14	I56574	heat shock protein 7	4.23e+00
24	83	12.9	641	14	I54542	heat shock protein 7	4.23e+00
25	83	12.9	641	7	A45871	heat shock protein H	4.23e+00
26	83	12.9	641	7	S35718	heat shock protein H	4.23e+00
27	83	12.9	642	7	JH0095	heat shock protein 7	4.23e+00
28	83	12.9	2647	13	A37098	gelation factor ABP-	4.23e+00
29	82	12.8	209	16	S48025	heat shock protein -	5.56e+00
30	82	12.8	405	15	A61181	homeotic protein HOX	5.56e+00
31	82	12.8	634	7	A25646	heat shock protein 7	5.56e+00
32	82	12.8	639	7	JC1391	heat shock protein 7	5.56e+00
33	82	12.8	645	12	I51129	heat shock protein H	5.56e+00
34	82	12.8	651	11	JC4786	heat-shock cognate p	5.56e+00
35	82	12.8	656	7	A48439	heat shock protein H	5.56e+00
36	82	12.8	814	7	S31211	collagen alpha 1(XIV	5.56e+00
37	82	12.8	874	4	QQBE15	BSLF1 protein - huma	5.56e+00
38	82	12.8	1747	7	A45974	collagen alpha 1(XIV	5.56e+00
39	81	12.6	214	7	A03309	heat shock-related 7	7.30e+00
40	81	12.6	641	7	JN0668	Heat-shock cognate p	7.30e+00
41	81	12.6	644	7	A45635	heat shock protein h	7.30e+00
42	81	12.6	654	7	S27004	heat shock protein 7	7.30e+00
43	80	12.4	171	11	PQ0475	pistil extensin-like	9.56e+00
44	80	12.4	645	7	S41372	heat shock protein -	9.56e+00
45	80	12.4	656	7	S51712	heat shock protein 7	9.56e+00

ALIGNMENTS

RESULT 1
ENTRY S58875 #type complete
TITLE cdn-2 protein - human
ORGANISM #formal_name Homo sapiens #common_name man
DATE 15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 01-Mar-1996
ACCESSIONS S58875
REFERENCE S58874
#authors Kiefer, M.C.; Brauer, M.J.; Powers, V.C.; Wu, J.J.; Umansky, S.R.; Tomei, L.D.; Barr, P.J.
#journal Nature (1995) 374:736-739
#title Modulation of apoptosis by the widely distributed Bcl-2 homologue Bak.
#accession S58875
##status preliminary; nucleic acid sequence not shown; translation not shown
##molecule_type DNA
##residues 1-211 ##label KIE
##cross-references EMBL:U16812
##note the nucleotide sequence was submitted to the EMBL Data Library, November 1994
SUMMARY #length 211 #molecular-weight 23411 #checksum 9485

Query Match 33.1%; Score 213; DB 13; Length 211;
Best Local Similarity 78.0%; Pred. No. 9.18e-20;
Matches 32; Conservative 2; Mismatches 6; Indels 1; Gaps 1;

```
Db      1 masgqgpgpprqecgepalpsaseeqvaqdteevfrsyvfy 41
          |||||||||||||| |||||||||||||| |  | :  |:
Qy      1 MASGQGPGPPRQECGKPALPSASEEQVAQDMEG-FSAATFF 40
```

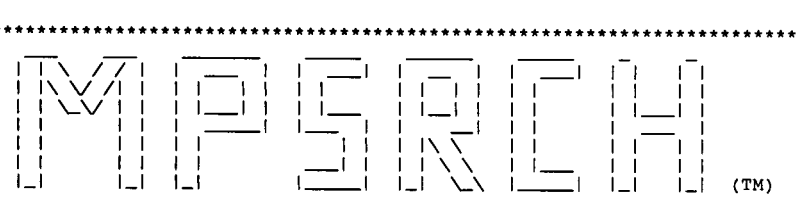
RESULT 2
ENTRY S58873 #type complete
TITLE Bak protein - human
ALTERNATE_NAMES bcl-2 homolog; cdn-1 protein
ORGANISM #formal_name Homo sapiens #common_name man
DATE 15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 06-Sep-1996
ACCESSIONS S58873; S58872; S58874
REFERENCE S58873
#authors Chittenden, T.; Harrington, E.A.; O'Connor, R.; Flemington, C.; Lutz, R.J.; Evan, G.I.; Guild, B.C.
#journal Nature (1995) 374:733-736
#title Induction of apoptosis by the Bcl-2 homologue Bak.

```
   #accession     S58873
      ##status        preliminary; nucleic acid sequence not shown
      ##molecule_type mRNA
      ##residues      1-211 ##label CHI
      ##cross-references EMBL:U23765
REFERENCE         S58872
   #authors       Farrow, S.N.; White, J.H.M.; Martinou, I.; Raven, T.; Pun,
                    K.T.; Grinham, C.J.; Martinou, J.C.; Brown, R.
   #journal       Nature (1995) 374:731-733
   #title         Cloning of a bcl-2 homologue by interaction with adenovirus
                    E1B 19K.
   #accession     S58872
      ##status        preliminary
      ##molecule_type mRNA
      ##residues      1-211 ##label FAR
      ##cross-references EMBL:X84213
REFERENCE         S58874
   #authors       Kiefer, M.C.; Brauer, M.J.; Powers, V.C.; Wu, J.J.; Umansky,
                    S.R.; Tomei, L.D.; Barr, P.J.
   #journal       Nature (1995) 374:736-739
   #title         Modulation of apoptosis by the widely distributed Bcl-2
                    homologue Bak.
   #accession     S58874
      ##status        preliminary
      ##molecule_type mRNA
      ##residues      1-211 ##label KIE
      ##cross-references EMBL:U16811
GENETICS
   #gene          GDB:BAK-LSB
      ##cross-references GDB:635887
SUMMARY           #length 211  #molecular-weight 23409  #checksum 801

  Query Match           33.1%;  Score 213;  DB 13;  Length 211;
  Best Local Similarity  78.0%;  Pred. No. 9.18e-20;
  Matches    32;  Conservative    2;  Mismatches   6;  Indels   1;  Gaps   1;

Db       1 masgqgpgpprqecgepalpsaseeqvaqdteevfrsyvfy 41
           ||||||||||||||| |||||||||||||| |  | :  |:
Qy       1 MASGQGPGPPRQECGKPALPSASEEQVAQDMEG-FSAATFF 40


RESULT    3
ENTRY             A40623     #type complete
TITLE             heat shock protein Y - Escherichia coli
ORGANISM          #formal_name Escherichia coli
DATE              21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change
                    15-Nov-1996
ACCESSIONS        A40623; A56688; S28460
REFERENCE         A40623
   #authors       Missiakas, D.; Georgopoulos, C.; Raina, S.
   #journal       J. Bacteriol. (1993) 175:2613-2624
   #title         The Escherichia coli heat shock gene htpY: mutational
                    analysis, cloning, sequencing, and transcriptional
                    regulation.
   #cross-references MUID:93239687
   #accession     A40623
      ##status        preliminary
      ##molecule_type DNA
      ##residues      1-196 ##label MIS
      ##cross-references NCBIN:130422; NCBIP:130423
      ##note          sequence extracted from NCBI backbone
REFERENCE         A56688
   #authors       James, R.; Dean, D.O.; Debbage, J.
   #journal       DNA Seq. (1993) 3:327-332
   #title         Five open reading frames upstream of the dnaK gene of
                    Escherichia coli.
   #accession     A56688
      ##status        preliminary
      ##molecule_type DNA
      ##residues      36-196 ##label JAM
      ##cross-references EMBL:X67700
GENETICS
   #gene          htpY; htgA
   #map_position 0 min
KEYWORDS          heat shock; stress-induced protein
SUMMARY           #length 196  #molecular-weight 21225  #checksum 5329

  Query Match           14.8%;  Score 95;  DB 9;  Length 196;
  Best Local Similarity  57.1%;  Pred. No. 1.35e-01;
  Matches    12;  Conservative    2;  Mismatches   7;  Indels   0;  Gaps   0;

Db      14 ppspaprskpcpstliaawvr 34
           || |: |  |||| | |:| |
Qy      53 PPLPTQRWSPCPSNLAAPWGR 73


RESULT    4
ENTRY             S37577     #type complete
TITLE             Shb protein - human
ORGANISM          #formal_name Homo sapiens #common_name man
DATE              06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change
                    06-Jan-1995
ACCESSIONS        S37577
REFERENCE         S37577
   #authors       Welsh, M.
   #submission    submitted to the EMBL Data Library, October 1993
   #accession     S37577
      ##status        preliminary
      ##molecule_type mRNA
      ##residues      1-596 ##label WEL
      ##cross-references EMBL:X75342
SUMMARY           #length 596  #molecular-weight 64525  #checksum 5363

  Query Match           13.8%;  Score 89;  DB 13;  Length 596;
  Best Local Similarity  33.3%;  Pred. No. 7.82e-01;
  Matches    13;  Conservative    9;  Mismatches  16;  Indels   1;  Gaps   1;

Db      48 rpggpppgpg-rrgtcaaalppewprrrtglprrgprpp 85
           |  | || |  |:: |:: |:: | |:  : |  | : |
Qy      48 RLKGRPPLPTQRWSPCPSNLAAPWGRWDGSSPSPGRHQP 86


RESULT    5
ENTRY             S09036     #type complete
TITLE             heat shock protein 70B' - human
ORGANISM          #formal_name Homo sapiens #common_name man
DATE              30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change
                    06-Sep-1996
ACCESSIONS        S09036
REFERENCE         S09036
   #authors       Leung, T.K.C.; Rajendran, M.Y.; Monfries, C.; Hall, C.; Lim,
                    L.
   #journal       Biochem. J. (1990) 267:125-132
   #title         The human heat-shock protein family. Expression of a novel
                    heat-inducible HSP70 (HSP70B') and isolation of its cDNA
                    and genomic DNA.
   #cross-references MUID:90226304
   #accession     S09036
      ##molecule_type DNA
      ##residues      1-643 ##label LEU
      ##cross-references EMBL:X51757
GENETICS
   #gene          GDB:HSPA6
      ##cross-references GDB:127963
   #map_position 1cen-1qter
CLASSIFICATION    #superfamily heat shock protein 70
SUMMARY           #length 643  #molecular-weight 70853  #checksum 313

  Query Match           13.4%;  Score 86;  DB 7;  Length 643;
  Best Local Similarity  26.8%;  Pred. No. 1.84e+00;
  Matches    11;  Conservative   12;  Mismatches  17;  Indels   1;  Gaps   1;

Db     263 grlrtacerakrtlssstqatleidslfegvdfytsitrar 303
           |  |  | ::    :|  | : :::: | :  |:|:|:| |
Qy       8 GPPRQECGKPALPSASEEQVAQDMEG-FSAATFFTTISRNR 47
```

```
RESULT    6
ENTRY             S18156     #type complete
TITLE             globulin 1 - eastern white pine
ORGANISM          #formal_name Pinus strobus #common_name eastern white pine
DATE              06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change
                    06-Jan-1995
ACCESSIONS        S18156
REFERENCE         S18156
  #authors        Rugh, C.L.; Kamalay, J.C.
  #submission     submitted to the EMBL Data Library, November 1991
  #description    Legumin mRNAs from Pinus strobus L.
  #accession      S18156
    ##status          preliminary
    ##molecule_type mRNA
    ##residues        1-488 ##label RUG
    ##cross-references EMBL:Z11486
SUMMARY           #length 488  #molecular-weight 54715  #checksum 3250

  Query Match           13.2%;  Score 85;  DB 10;  Length 488;
  Best Local Similarity  36.1%;  Pred. No. 2.43e+00;
  Matches    13;  Conservative    7;  Mismatches  14;  Indels    2;  Gaps    2;

Db    248 ffarvergrrlslpaprsrsrsrspyagr-qrqwgr 282
          ||: : | |||   :| :  | :|| ::     |||
Qy     39 FFTTISRNRRLKGRPPLPTQR-WSPCPSNLAAPWGR 73


RESULT    7
ENTRY             A36333     #type complete
TITLE             heat shock cognate protein 70 - fruit fly (Drosophila
                    melanogaster)
ORGANISM          #formal_name Drosophila melanogaster
DATE              25-Jan-1991 #sequence_revision 25-Jan-1991 #text_change
                    12-Apr-1995
ACCESSIONS        A36333
REFERENCE         A36333
  #authors        Perkins, L.A.; Doctor, J.S.; Zhang, K.; Stinson, L.;
                    Perrimon, N.; Craig, E.A.
  #journal        Mol. Cell. Biol. (1990) 10:3232-3238
  #title          Molecular and developmental characterization of the heat
                    shock cognate 4 gene of Drosophila melanogaster.
  #cross-references MUID:90258915
  #accession      A36333
    ##status          preliminary
    ##molecule_type DNA
    ##residues        1-651 ##label PER
    ##cross-references GB:M36114
CLASSIFICATION    #superfamily heat shock protein 70
SUMMARY           #length 651  #molecular-weight 71059  #checksum 805

  Query Match           13.2%;  Score 85;  DB 7;  Length 651;
  Best Local Similarity  26.3%;  Pred. No. 2.43e+00;
  Matches    10;  Conservative   13;  Mismatches  14;  Indels    1;  Gaps    1;

Db    264 rtacerakrtlssstqasieidslfegtdfytsitrar 301
          |  | ::    :|  | : :::: | :: |:|:|:| |
Qy     11 RQECGKPALPSASEEQVAQDMEG-FSAATFFTTISRNR 47


RESULT    8
ENTRY             A48872     #type fragment
TITLE             70k heat shock cognate protein aginactin - slime mold
                    (Dictyostelium discoideum) (fragment)
ALTERNATE_NAMES   F-actin capping protein aginactin
ORGANISM          #formal_name Dictyostelium discoideum
DATE              19-May-1995 #sequence_revision 19-May-1995 #text_change
                    03-May-1996
ACCESSIONS        A48872
REFERENCE         A48872
  #authors        Eddy, R.J.; Sauterer, R.A.; Condeelis, J.S.
  #journal        J. Biol. Chem. (1993) 268:23267-23274
  #title          Aginactin, an agonist-regulated F-actin capping activity is
                    associated with an Hsc70 in Dictyostelium.
  #accession      A48872
    ##status          preliminary
    ##molecule_type mRNA
    ##residues        1-636 ##label EDD
    ##cross-references GB:L22736
    ##note            authors translated the codon ACA for residue 10 as Pro,
                      GCT for residue 60 as Val, GGT for residue 350 as Thr,
                      and TTA for residue 575 as Pro
CLASSIFICATION    #superfamily heat shock protein 70
SUMMARY           #length 636  #checksum 2081

  Query Match           13.1%;  Score 84;  DB 7;  Length 636;
  Best Local Similarity  26.3%;  Pred. No. 3.21e+00;
  Matches    10;  Conservative   12;  Mismatches  15;  Indels    1;  Gaps    1;

Db    258 rtacerakrtlsssaqasieidslfegidfytsitrar 295
          |  | ::    :|  | : :::: | :  |:|:|:| |
Qy     11 RQECGKPALPSASEEQVAQDMEG-FSAATFFTTISRNR 47


RESULT    9
ENTRY             S20139     #type complete
TITLE             heat shock protein SSA2 - yeast (Saccharomyces cerevisiae)
ALTERNATE_NAMES   heat shock protein YG102; protein L0971; protein YLL024c
ORGANISM          #formal_name Saccharomyces cerevisiae
DATE              07-May-1993 #sequence_revision 07-May-1993 #text_change
                    06-Sep-1996
ACCESSIONS        S20139; S64772; S64775; S69383
REFERENCE         S20139
  #authors        Slater, M.R.; Craig, E.A.
  #journal        Nucleic Acids Res. (1989) 17:805-806
  #title          The SSA1 and SSA2 genes of the yeast Saccharomyces
                    cerevisiae.
  #cross-references MUID:89128457
  #accession      S20139
    ##status          translation not shown
    ##molecule_type DNA
    ##residues        1-639 ##label SLA
    ##cross-references EMBL:X12927
REFERENCE         S64761
  #authors        Goffeau, A.; Purnelle, B.
  #submission     submitted to the Protein Sequence Database, May 1996
  #accession      S64772
    ##molecule_type DNA
    ##residues        1-639 ##label GOF
    ##cross-references EMBL:Z73129
    ##experimental_source strain S288C
REFERENCE         S64775
  #authors        Duesterhoeft, A.; Floeth, M.; Heuss-Neitzel, D.; Hilbert, H.;
                    Moestl, D.
  #submission     submitted to the Protein Sequence Database, May 1996
  #accession      S64775
    ##molecule_type DNA
    ##residues        72-639 ##label DUE
    ##cross-references EMBL:Z73129
    ##experimental_source strain S288C
REFERENCE         S69380
  #authors        Purnelle, B.; Goffeau, A.
  #submission     submitted to the EMBL Data Library, April 1996
  #description    The sequence of 32 kb on the left arm of yeast chromosome XII
                    reveals 14 open reading frames among which HSP104, SSA2,
                    SPA2, KNS1, DPS1/APS, SDC25, a new member of the
                    seripauperins family and a new ABC transporter homologous
                    to the human multidrug resistance protein.
  #accession      S69383
    ##molecule_type DNA
    ##residues        1-639 ##label PUR
    ##cross-references EMBL:X97560
GENETICS
  #gene           LISTA:SSA2
  #map_position   12L
CLASSIFICATION    #superfamily heat shock protein 70
```

```
KEYWORDS          molecular chaperone
SUMMARY           #length 639  #molecular-weight 69469  #checksum 5309

  Query Match           13.1%;  Score 84;  DB 7;  Length 639;
  Best Local Similarity  26.3%;  Pred. No. 3.21e+00;
  Matches    10;  Conservative   12;  Mismatches  15;  Indels   1;  Gaps   1;

Db    261 rtacerakrtlsssaqtsveidslfegidfytsitrar 298
          |  | ::    :|  | : :::: | :  |:|:|:| |
Qy     11 RQECGKPALPSASEEQVAQDMEG-FSAATFFTTISRNR 47


RESULT   10
ENTRY             S37394     #type complete
TITLE             heat shock cognate protein 70 - slime mold (Dictyostelium
                    discoideum)
ORGANISM          #formal_name Dictyostelium discoideum
DATE              31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change
                    12-Apr-1995
ACCESSIONS        S37394
REFERENCE         S37394
   #authors       Haus, U.; Trommler, P.; Fisher, P.R.; Hartmann, H.;
                    Lottspeich, F.; Noegel, A.A.; Schleicher, M.
   #journal       EMBO J. (1993) 12:3763-3771
   #title         The heat shock cognate protein from Dictyostelium affects
                    actin polymerization through interaction with the
                    actin-binding protein cap32/34.
   #accession     S37394
      ##molecule_type mRNA
      ##residues      1-640 ##label HAU
      ##cross-references EMBL:X75263
GENETICS
   #gene          hsc70
CLASSIFICATION    #superfamily heat shock protein 70
SUMMARY           #length 640  #molecular-weight 70499  #checksum 4857

  Query Match           13.1%;  Score 84;  DB 7;  Length 640;
  Best Local Similarity  26.3%;  Pred. No. 3.21e+00;
  Matches    10;  Conservative   12;  Mismatches  15;  Indels   1;  Gaps   1;

Db    262 rtacerakrtlsssaqasieidslfegidfytsitrar 299
          |  | ::    :|  | : :::: | :  |:|:|:| |
Qy     11 RQECGKPALPSASEEQVAQDMEG-FSAATFFTTISRNR 47


RESULT   11
ENTRY             HHBYA1     #type complete
TITLE             heat shock protein SSA1 - yeast (Saccharomyces cerevisiae)
ALTERNATE_NAMES   heat shock protein 70-related protein SSA1; heat shock
                    protein YG100; protein YAL005c
ORGANISM          #formal_name Saccharomyces cerevisiae
DATE              30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change
                    15-Nov-1996
ACCESSIONS        S43449; S25438; S42164; S25231; S40897
REFERENCE         S43441
   #authors       Clark, M.W.; Keng, T.; Storms, R.K.; Zhong, W.; Fortin, N.;
                    Zeng, B.; Delaney, S.; Ouellette, B.F.F.; Barton, A.B.;
                    Kaback, D.B.; Bussey, H.
   #journal       Yeast (1994) 10:535-541
   #title         Sequencing of chromosome I of Saccharomyces cerevisiae:
                    analysis of the 42 kbp SPO7-CENI-CDC15 region.
   #accession     S43449
      ##status        nucleic acid sequence not shown; translation not shown
      ##molecule_type DNA
      ##residues      1-642 ##label CLA
      ##cross-references EMBL:L22015
      ##note          the nucleotide sequence was submitted to the EMBL Data
                       Library, November 1993
REFERENCE         S20139
   #authors       Slater, M.R.; Craig, E.A.
   #journal       Nucleic Acids Res. (1989) 17:805-806
   #title         The SSA1 and SSA2 genes of the yeast Saccharomyces
                    cerevisiae.
   #cross-references MUID:89128457
   #accession     S25438
      ##status        translation not shown
      ##molecule_type DNA
      ##residues      1-642 ##label SLA
      ##cross-references EMBL:X12926
      ##note          this sequence has been revised in reference S42164
REFERENCE         S42164
   #authors       Slater, M.R.
   #submission    submitted to the EMBL Data Library, June 1993
   #accession     S42164
      ##molecule_type DNA
      ##residues      1-207,'S',209-417,'P',419-421,'S',423-642 ##label SL2
      ##cross-references EMBL:X12926
      ##note          this is a revision to the sequence from reference S20139
GENETICS
   #gene          LISTA:SSA1; YG100
   #map_position 1L
CLASSIFICATION    #superfamily heat shock protein 70
KEYWORDS          heat shock; molecular chaperone; stress-induced protein;
                    transmembrane protein
FEATURE
   193-209            #domain transmembrane #status predicted #label TMM
SUMMARY           #length 642  #molecular-weight 69767  #checksum 6411

  Query Match           13.1%;  Score 84;  DB 3;  Length 642;
  Best Local Similarity  26.3%;  Pred. No. 3.21e+00;
  Matches    10;  Conservative   12;  Mismatches  15;  Indels   1;  Gaps   1;

Db    261 rtacerakrtlsssaqtsveidslfegidfytsitrar 298
          |  | ::    :|  | : :::: | :  |:|:|:| |
Qy     11 RQECGKPALPSASEEQVAQDMEG-FSAATFFTTISRNR 47


RESULT   12
ENTRY             S25585     #type complete
TITLE             heat shock protein 70 - pig
ORGANISM          #formal_name Sus scrofa domestica #common_name domestic pig
DATE              20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change
                    03-Aug-1995
ACCESSIONS        S34625; S25585
REFERENCE         S34625
   #authors       Dezeure, F.; Vaiman, M.; Chardon, P.
   #journal       Biochim. Biophys. Acta (1993) 1174:17-26
   #title         Characterization of a polymorphic heat shock protein 70 gene
                    in swine outside the SLA major histocompatibility complex.
   #accession     S34625
      ##status        preliminary
      ##molecule_type mRNA
      ##residues      1-643 ##label DEZ
      ##cross-references EMBL:X68213
CLASSIFICATION    #superfamily heat shock protein 70
SUMMARY           #length 643  #molecular-weight 71109  #checksum 9470

  Query Match           13.1%;  Score 84;  DB 7;  Length 643;
  Best Local Similarity  26.3%;  Pred. No. 3.21e+00;
  Matches    10;  Conservative   12;  Mismatches  15;  Indels   1;  Gaps   1;

Db    266 rtacerakrtlssstqatleidslfegvdfytsitrar 303
          |  | ::    :|  | : :::: | :  |:|:|:| |
Qy     11 RQECGKPALPSASEEQVAQDMEG-FSAATFFTTISRNR 47


RESULT   13
ENTRY             HHXL70     #type complete
TITLE             heat shock protein 70 - African clawed frog
ALTERNATE_NAMES   heat shock protein X16
ORGANISM          #formal_name Xenopus laevis #common_name African clawed frog
DATE              28-May-1986 #sequence_revision 28-May-1986 #text_change
                    15-Nov-1996
ACCESSIONS        A03310; A22175
REFERENCE         A90993
   #authors       Bienz, M.
```

```
   #journal      EMBO J. (1984) 3:2477-2483
   #title        Xenopus hsp 70 genes are constitutively expressed in injected
                   oocytes.
   #cross-references MUID:85076567
   #accession    A03310
      ##molecule_type DNA
      ##residues      1-647 ##label BIE
REFERENCE        A22175
   #authors      Bienz, M.
   #journal      Proc. Natl. Acad. Sci. U.S.A. (1984) 81:3138-3142
   #title        Developmental control of the heat shock response in Xenopus.
   #cross-references MUID:84221917
   #accession    A22175
      ##molecule_type mRNA
      ##residues      81-120 ##label BI2
COMMENT     This protein is expressed constitutively in oocytes, disappears
              after fertilization, and is induced by heat shock in somatic
              cells from the gastrula stage onward.
CLASSIFICATION   #superfamily heat shock protein 70
KEYWORDS         heat shock; molecular chaperone; stress-induced protein
SUMMARY          #length 647  #molecular-weight 70915  #checksum 3965

  Query Match           13.1%;  Score 84;  DB 3;  Length 647;
  Best Local Similarity  26.3%;  Pred. No. 3.21e+00;
  Matches    10;  Conservative   12;  Mismatches  15;  Indels   1;  Gaps   1;

Db    265 rtacdrakrtlssssqasieidslfegidfytaitrar 302
          |  | ::    :|  | : :::: | :  |:|:|:| |
Qy     11 RQECGKPALPSASEEQVAQDMEG-FSAATFFTTISRNR 47


RESULT   14
ENTRY            S18349      #type complete
TITLE            heat shock protein 70 - carrot
ORGANISM         #formal_name Daucus carota #common_name carrot
DATE             04-Dec-1992 #sequence_revision 04-Dec-1992 #text_change
                   12-Apr-1995
ACCESSIONS       S18349; S15514
REFERENCE        S18349
   #authors      Lin, X.; Chern, M.; Zimmerman, J.L.
   #journal      Plant Mol. Biol. (1991) 17:1245-1249
   #title        Cloning and characterization of a carrot hsp70 gene.
   #cross-references MUID:92032789
   #accession    S18349
      ##molecule_type DNA
      ##residues      1-655 ##label LIN
      ##cross-references EMBL:X60088
CLASSIFICATION   #superfamily heat shock protein 70
SUMMARY          #length 655  #molecular-weight 72051  #checksum 2153

  Query Match           13.1%;  Score 84;  DB 7;  Length 655;
  Best Local Similarity  30.6%;  Pred. No. 3.21e+00;
  Matches    11;  Conservative   14;  Mismatches   9;  Indels   2;  Gaps   2;

Db    272 erakrtlsssta-qttieidslyegvdfyttitrar 306
          | :| :|:|::  | : :::: : :  |:|||:| |
Qy     13 ECGKPALPSASEEQVAQDMEG-FSAATFFTTISRNR 47


RESULT   15
ENTRY            S26369      #type complete
TITLE            genome polyprotein - eastern equine encephalomyelitis virus
CONTAINS         nonstructural protein NS1; nonstructural protein NS2;
                   nonstructural protein NS3
ORGANISM         #formal_name eastern equine encephalomyelitis virus
DATE             06-Jan-1994 #sequence_revision 01-Nov-1996 #text_change
                   01-Nov-1996
ACCESSIONS       S26369
REFERENCE        S26369
   #authors      Volchkov, V.E.; Volchkova, V.A.; Netesov, S.V.
   #journal      Mol. Gen. Mikrobiol. Virusol. (1991) 5:8-15
   #title        Complete nucleotide sequence of the eastern equine
                   encephalomyelitis virus genome.
   #accession    S26369
      ##molecule_type mRNA
      ##residues      1-1878 ##label VOL
      ##cross-references EMBL:X63135
      ##note          sequence could not be checked because of bad print in
                       paper
KEYWORDS         nonstructural protein; polyprotein
FEATURE
   1-532              #product nonstructural protein NS1 #status predicted
                        #label NS1\
   533-1326           #product nonstructural protein NS2 #status predicted
                        #label NS2\
   1327-1878          #product nonstructural protein NS3 #status predicted
                        #label NS3
SUMMARY          #length 1878  #molecular-weight 208582  #checksum 5340

  Query Match           13.1%;  Score 84;  DB 8;  Length 1878;
  Best Local Similarity  33.9%;  Pred. No. 3.21e+00;
  Matches    19;  Conservative   14;  Mismatches  19;  Indels   4;  Gaps   4;

Db   1774 psppsqd-srpstpsasashtfvdlitfdsvaeiledfsrspfqflseikpipapr 1828
          |:|| |: ::|: |||| :   |:  | | | ::  :||:  : |    |:|: |
Qy      7 PGPPRQECGKPALPSASEEQVAQDMEGF-SAATFFTTISRNR-R-LKGRPPLPTQR 59


Search completed: Wed Aug 20 11:09:37 1997
Job time : 36 secs.
```



MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Wed Aug 20 11:09:54 1997; MasPar time 3.60 Seconds
517.994 Million cell updates/sec

Tabular output not generated.

Title: >US-08-320-157-21
Description: (1-88) from US08320157.pep
Perfect Score: 643
Sequence: 1 MASGQGPGPPRQECGKPALP.....APWGRWDGSSPSPGRHQPAL 88

Scoring table: PAM 150
Gap 11

Searched: 59021 seqs, 21210388 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: swiss-prot34
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11

Statistics: Mean 38.620; Variance 68.467; scale 0.564

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description	Pred. No.
1	95	14.8	196	5	HTGA_ECOLI	HEAT SHOCK PROTEIN HT	1.66e-02
2	86	13.4	643	5	HS76_HUMAN	HEAT SHOCK 70 KD PROT	3.58e-01
3	85	13.2	651	5	HS7D_DROME	HEAT SHOCK 70 KD PROT	4.97e-01
4	84	13.1	379	5	HS7X_PIG	HEAT SHOCK 70 KD PROT	6.89e-01
5	84	13.1	631	5	HS73_BOVIN	HEAT SHOCK 70 KD PROT	6.89e-01
6	84	13.1	638	5	HS72_YEAST	HEAT SHOCK PROTEIN SS	6.89e-01
7	84	13.1	640	5	HS7C_DICDI	HEAT SHOCK COGNATE PR	6.89e-01
8	84	13.1	641	5	HS71_YEAST	HEAT SHOCK PROTEIN SS	6.89e-01
9	84	13.1	643	5	HS76_PIG	HEAT SHOCK 70 KD PROT	6.89e-01
10	84	13.1	647	5	HS70_XENLA	HEAT SHOCK 70 KD PROT	6.89e-01
11	84	13.1	649	5	HS70_BLAEM	HEAT SHOCK 70 KD PROT	6.89e-01
12	84	13.1	655	5	HS70_DAUCA	HEAT SHOCK 70 KD PROT	6.89e-01
13	84	13.1	2142	1	BAT2_HUMAN	LARGE PROLINE-RICH PR	6.89e-01
14	83	12.9	284	4	GOLI_DROME	GOLIATH PROTEIN (G1 P	9.52e-01
15	83	12.9	462	4	FUCO_RAT	ALPHA-L-FUCOSIDASE PR	9.52e-01
16	83	12.9	641	5	HS71_HUMAN	HEAT SHOCK 70 KD PROT	9.52e-01
17	83	12.9	641	5	HS71_PIG	HEAT SHOCK 70 KD PROT	9.52e-01
18	83	12.9	641	5	HS71_RAT	HEAT SHOCK 70 KD PROT	9.52e-01
19	83	12.9	642	5	HS71_MOUSE	HEAT SHOCK 70 KD PROT	9.52e-01
20	83	12.9	642	5	HS72_PICAN	HEAT SHOCK PROTEIN 70	9.52e-01
21	83	12.9	2647	1	ABP2_HUMAN	ENDOTHELIAL ACTIN-BIN	9.52e-01
22	82	12.8	634	5	HS70_CHICK	HEAT SHOCK 70 KD PROT	1.31e+00
23	82	12.8	639	5	HS74_PARLI	HEAT SHOCK 70 KD PROT	1.31e+00
24	82	12.8	787	8	RIR1_HSVBC	RIBONUCLEOSIDE-DIPHOS	1.31e+00
25	82	12.8	874	10	UL52_EBV	HELICASE/PRIMASE COMP	1.31e+00
26	82	12.8	1888	2	CA1E_CHICK	COLLAGEN ALPHA 1(XIV)	1.31e+00
27	81	12.6	214	5	HS7A_DROSI	HEAT SHOCK 70 KD PROT	1.80e+00
28	81	12.6	322	5	HS70_ONCVO	HEAT SHOCK 70 KD PROT	1.80e+00
29	81	12.6	641	5	HS7A_DROME	HEAT SHOCK 70 KD PROT	1.80e+00
30	81	12.6	644	5	HS70_BRUMA	HEAT SHOCK 70 KD PROT	1.80e+00
31	81	12.6	654	5	HS70_HYDMA	HEAT SHOCK 70 KD PROT	1.80e+00
32	80	12.4	641	5	HS72_DROME	MAJOR HEAT SHOCK 70 K	2.47e+00
33	80	12.4	643	5	HS71_DROME	MAJOR HEAT SHOCK 70 K	2.47e+00
34	80	12.4	643	5	HS71_SCHPO	PROBABLE HEAT SHOCK P	2.47e+00
35	80	12.4	644	5	HS71_HANPO	HEAT-SHOCK PROTEIN 70	2.47e+00
36	80	12.4	650	5	HS71_LYCES	HEAT SHOCK COGNATE 70	2.47e+00
37	80	12.4	655	5	HS74_CANAL	HEAT SHOCK PROTEIN SS	2.47e+00
38	80	12.4	1403	6	NKCR_HUMAN	NK-TUMOR RECOGNITION	2.47e+00
39	80	12.4	1466	6	NKCR_MOUSE	NK-TUMOR RECOGNITION	2.47e+00
40	80	12.4	1822	11	YM68_CAEEL	HYPOTHETICAL HELICASE	2.47e+00
41	79	12.3	95	10	VE3_BPV2	PROBABLE E3 PROTEIN.	3.38e+00
42	79	12.3	408	11	YHGL_ECOLI	HYPOTHETICAL 45.1 KD	3.38e+00
43	79	12.3	633	5	HS72_MOUSE	HEAT SHOCK-RELATED 70	3.38e+00
44	79	12.3	639	5	HS72_HUMAN	HEAT SHOCK-RELATED 70	3.38e+00
45	79	12.3	651	5	HS7C_PETHY	HEAT SHOCK COGNATE 70	3.38e+00

ALIGNMENTS

```
RESULT    1
ID   HTGA_ECOLI     STANDARD;      PRT;   196 AA.
AC   P28697;
DT   01-DEC-1992 (REL. 24, CREATED)
DT   01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)
DT   01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)
DE   HEAT SHOCK PROTEIN HTGA (HEAT SHOCK PROTEIN HTPY).
GN   HTGA OR HTPY.
OS   ESCHERICHIA COLI.
OC   PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
OC   ENTEROBACTERIACEAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 94003405.
RA   JAMES R., DEAN D.O., DEBBAGE J.;
RL   DNA SEQ. 3:327-332(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12 / W3110;
RX   MEDLINE; 93239687.
RA   MISSIAKAS D., GEORGOPOULOS C., RAINA S.;
RL   J. BACTERIOL. 175:2613-2624(1993).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12;
RX   MEDLINE; 92334977.
RA   YURA T., MORI H., NAGAI H., NAGATA T., ISHIHAMA A., FUJITA N.,
RA   ISONO K., MIZOBUCHI K., NAKATA A.;
RL   NUCLEIC ACIDS RES. 20:3305-3308(1992).
CC   -!- FUNCTION: NOT KNOWN, REQUIRED FOR HIGH-TEMPERATURE GROWTH OF
CC       E.COLI. IT IS POSSIBLE THAT HTGA PROTEIN SOMEHOW REGULATES EITHER
CC       TRANSCRIPTION OF THE RPOH GENE OR THE ACTIVITY OF ITS GENE PRODUCT
CC       SIGMA-32. PLAYS A ROLE OPPOSITE THAT OF DNAK, DNAJ, AND GRPE IN
CC       TERMS OF HEAT SHOCK REGULATION, ANTAGONIZING THE NEGATIVE OF
CC       THESE AND THUS FINE-TUNING THE HEAT SHOCK RESPONSE.
CC   -!- INDUCTION: INDUCED AT HIGH TEMPERATURES.
CC   -!- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-36 IS THE INITIATOR.
CC   -!- CAUTION: REF.3 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A
CC       FRAMESHIFT.
DR   EMBL; X67700; G41757; ALT_INIT.
DR   EMBL; L03720; G290451; -.
DR   EMBL; D10483; -; NOT_ANNOTATED_CDS.
DR   PIR; S28460; S28460.
DR   PIR; A40623; A40623.
DR   ECOGENE; EG11509; HTGA.
KW   HEAT SHOCK.
```

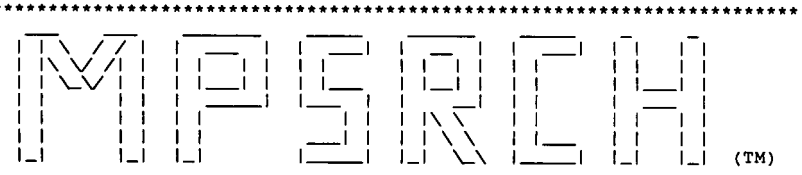
```
FT   DOMAIN       57     65       POLY-SER.
SQ   SEQUENCE   196 AA;  21225 MW;  D9E3CEC8 CRC32;

  Query Match           14.8%;  Score 95;  DB 5;  Length 196;
  Best Local Similarity  57.1%;  Pred. No. 1.66e-02;
  Matches    12;  Conservative    2;  Mismatches   7;  Indels   0;  Gaps   0;

Db     14 ppspaprskpcpstliaawvr 34
          || |: |  |||| | |:| |
Qy     53 PPLPTQRWSPCPSNLAAPWGR 73


RESULT    2
ID   HS76_HUMAN     STANDARD;      PRT;   643 AA.
AC   P17066;
DT   01-AUG-1990 (REL. 15, CREATED)
DT   01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)
DT   01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)
DE   HEAT SHOCK 70 KD PROTEIN 6 (HEAT SHOCK 70 KD PROTEIN B').
GN   HSPA6 OR HSP70B'.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LUNG;
RX   MEDLINE; 90226304.
RA   LEUNG T.K.C., RAJENDRAN M.Y., MONFRIES C., HALL C., LIM L.;
RL   BIOCHEM. J. 267:125-132(1990).
RN   [2]
RP   SEQUENCE OF 1-250 FROM N.A.
RX   MEDLINE; 92128997.
RA   LEUNG T.K.C., HALL C., RAJENDRAN M., SPURR N.K., LIM L.;
RL   GENOMICS 12:74-79(1992).
CC   -!- INDUCTION: ONLY AT HIGHER TEMPERATURES, AND NO BASAL EXPRESSION.
CC   -!- SIMILARITY: BELONGS TO HEAT SHOCK PROTEIN 70 FAMILY.
DR   EMBL; X51757; G35222; -.
DR   EMBL; X51758; G35224; -.
DR   EMBL; S78631; -; NOT_ANNOTATED_CDS.
DR   PIR; S09036; S09036.
DR   HSSP; P19120; 1NGA.
DR   HSC-2DPAGE; P17066; HUMAN.
DR   MIM; 140555; -.
DR   PROSITE; PS00297; HSP70_1.
DR   PROSITE; PS00329; HSP70_2.
DR   PROSITE; PS01036; HSP70_3.
KW   ATP-BINDING; HEAT SHOCK; MULTIGENE FAMILY.
SQ   SEQUENCE   643 AA;  70854 MW;  D7C64887 CRC32;

  Query Match           13.4%;  Score 86;  DB 5;  Length 643;
  Best Local Similarity  26.8%;  Pred. No. 3.58e-01;
  Matches    11;  Conservative   12;  Mismatches  17;  Indels   1;  Gaps   1;

Db    263 grlrtacerakrtlssstqatleidslfegvdfytsitrar 303
          |  |  | ::    :|  | : :::: | :  |:|:|:| |
Qy      8 GPPRQECGKPALPSASEEQVAQDMEG-FSAATFFTTISRNR 47


RESULT    3
ID   HS7D_DROME     STANDARD;      PRT;   651 AA.
AC   P11147;
DT   01-JUL-1989 (REL. 11, CREATED)
DT   01-APR-1993 (REL. 25, LAST SEQUENCE UPDATE)
DT   01-JUN-1994 (REL. 29, LAST ANNOTATION UPDATE)
DE   HEAT SHOCK 70 KD PROTEIN COGNATE 4 (HEAT SHOCK 70 KD PROTEIN 88E).
GN   HSC4 OR HSC70-4.
OS   DROSOPHILA MELANOGASTER (FRUIT FLY).
OC   EUKARYOTA; METAZOA; ARTHROPODA; INSECTA; DIPTERA.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 93292982.
RA   RUBIN D.M., MEHTA A., ZHU J., SHOHAM S., CHEN X., WELLS Q.,
RA   PALTER K.B.;
RL   GENE 128:155-163(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 90258915.
RA   PERKINS L.A., DOCTOR J.S., ZHANG K., STINSON L., PERRIMON N.,
RA   CRAIG E.A.;
RL   MOL. CELL. BIOL. 10:3232-3238(1990).
RN   [3]
RP   SEQUENCE OF 1-104 FROM N.A.
RX   MEDLINE; 84005511.
RA   CRAIG E.A., INGOLIA T.D., MANSEAU L.J.;
RL   DEV. BIOL. 99:418-426(1983).
RN   [4]
RP   SEQUENCE OF 580-592.
RC   STRAIN=VALLECAS; TISSUE=WING IMAGINAL DISCS;
RX   MEDLINE; 93272852.
RA   SANTAREN J.F., VAN DAMME J., PUYPE M., VANDEKERCKHOVE J.,
RA   GARCIA-BELLIDO A.;
RL   EXP. CELL RES. 206:220-226(1993).
CC   -!- DEVELOPMENTAL STAGE: HEAT SHOCK COGNATE PROTEINS ARE EXPRESSED
CC       CONSTITUTIVELY DURING NORMAL DEVELOPMENT.
CC   -!- SUBCELLULAR LOCATION: LOCALISED TO A MESHWORK OF CYTOPLASMIC
CC       FIBRES AROUND THE NUCLEUS. TRANSLOCATES TO THE NUCLEUS AFTER
CC       THERMAL STRESS.
CC   -!- SIMILARITY: BELONGS TO HEAT SHOCK PROTEIN 70 FAMILY.
DR   EMBL; L01500; G157661; -.
DR   EMBL; M36114; G157665; -.
DR   EMBL; J02569; G157678; ALT_SEQ.
DR   PIR; A36333; A36333.
DR   HSSP; P19120; 1ATR.
DR   FLYBASE; FBGN0001219; HSC70-4.
DR   PROSITE; PS00297; HSP70_1.
DR   PROSITE; PS00329; HSP70_2.
DR   PROSITE; PS01036; HSP70_3.
KW   ATP-BINDING; HEAT SHOCK; NUCLEAR PROTEIN; MULTIGENE FAMILY.
FT   CONFLICT    167    167       L -> P (IN REF. 2).
FT   CONFLICT    170    170       L -> P (IN REF. 2).
FT   CONFLICT    625    625       P -> G (IN REF. 2).
SQ   SEQUENCE   651 AA;  71131 MW;  75CC8721 CRC32;

  Query Match           13.2%;  Score 85;  DB 5;  Length 651;
  Best Local Similarity  26.3%;  Pred. No. 4.97e-01;
  Matches    10;  Conservative   13;  Mismatches  14;  Indels   1;  Gaps   1;

Db    264 rtacerakrtlssstqasieidslfegtdfytsitrar 301
          |  | ::    :|  | : :::: | :: |:|:|:| |
Qy     11 RQECGKPALPSASEEQVAQDMEG-FSAATFFTTISRNR 47


RESULT    4
ID   HS7X_PIG       STANDARD;      PRT;   379 AA.
AC   P34934;
DT   01-FEB-1994 (REL. 28, CREATED)
DT   01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)
DT   01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)
DE   HEAT SHOCK 70 KD PROTEIN (FRAGMENT).
OS   SUS SCROFA (PIG).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; ARTIODACTYLA.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=KIDNEY;
RX   MEDLINE; 90371455.
RA   BUCHMAN T.G., CABIN D.E., VICKERS S.;
RL   SURGERY 108:559-566(1990).
CC   -!- SIMILARITY: BELONGS TO HEAT SHOCK PROTEIN 70 FAMILY.
DR   EMBL; M29506; G164495; -.
DR   HSSP; P19120; 1NGA.
DR   PROSITE; PS00297; HSP70_1.
DR   PROSITE; PS00329; HSP70_2.
DR   PROSITE; PS01036; HSP70_3.
KW   ATP-BINDING; HEAT SHOCK; MULTIGENE FAMILY.
FT   NON_TER       1      1
```

```
SQ   SEQUENCE   379 AA;  42289 MW;  9E26ABE8 CRC32;

  Query Match           13.1%;  Score 84;  DB 5;  Length 379;
  Best Local Similarity  26.3%;  Pred. No. 6.89e-01;
  Matches    10;  Conservative   12;  Mismatches  15;  Indels   1;  Gaps   1;

Db      4 rtacerakrtlssstqatleidslfegvdfytsitrar 41
          |  | ::    :|  | : :::: | :  |:|:|:| |
Qy     11 RQECGKPALPSASEEQVAQDMEG-FSAATFFTTISRNR 47


RESULT    5
ID   HS73_BOVIN     STANDARD;      PRT;   631 AA.
AC   P34933;
DT   01-FEB-1994 (REL. 28, CREATED)
DT   01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   HEAT SHOCK 70 KD PROTEIN 3.
GN   HSP70-3.
OS   BOS TAURUS (BOVINE).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; ARTIODACTYLA.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   GROSZ M.D., MASSEY V.K., SKOW L.C.;
RL   SUBMITTED (XXX-1992) TO EMBL/GENBANK/DDBJ DATA BANKS.
CC   -!- FUNCTION: IN COOPERATION WITH OTHER CHAPERONES, HSP70S STABILIZE
CC       PREEXISTENT PROTEINS AGAINST AGGREGATION AND MEDIATE THE FOLDING
CC       OF NEWLY TRANSLATED POLYPEPTIDES IN THE CYTOSOL AS WELL AS WITHIN
CC       ORGANELLES. THESE CHAPERONES PARTICIPATE IN ALL THESE PROCESSES
CC       THROUGH THEIR ABILITY TO RECOGNIZE NONNATIVE CONFORMATIONS OF
CC       OTHER PROTEINS. THEY BIND EXTENDED PEPTIDE SEGMENTS WITH A NET
CC       HYDROPHOBIC CHARACTER EXPOSED BY POLYPEPTIDES DURING TRANSLATION
CC       AND MEMBRANE TRANSLOCATION, OR FOLLOWING STRESS-INDUCED DAMAGE.
CC   -!- SIMILARITY: BELONGS TO HEAT SHOCK PROTEIN 70 FAMILY.
DR   EMBL; L10428; G163161; -.
DR   HSSP; P19120; 1NGJ.
DR   PROSITE; PS00297; HSP70_1.
DR   PROSITE; PS00329; HSP70_2.
DR   PROSITE; PS01036; HSP70_3.
KW   ATP-BINDING; CHAPERONE; HEAT SHOCK; MULTIGENE FAMILY.
SQ   SEQUENCE   631 AA;  69199 MW;  C7B8B580 CRC32;

  Query Match           13.1%;  Score 84;  DB 5;  Length 631;
  Best Local Similarity  23.3%;  Pred. No. 6.89e-01;
  Matches    10;  Conservative   15;  Mismatches  17;  Indels   1;  Gaps   1;

Db    257 aparlrtacerakrtlssstqasieidslyegvdfytsitrar 299
          :|:  |  | ::    :|  | : :::: : :  |:|:|:| |
Qy      6 GPGPPRQECGKPALPSASEEQVAQDMEG-FSAATFFTTISRNR 47


RESULT    6
ID   HS72_YEAST     STANDARD;      PRT;   638 AA.
AC   P10592;
DT   01-JUL-1989 (REL. 11, CREATED)
DT   01-OCT-1994 (REL. 30, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   HEAT SHOCK PROTEIN SSA2.
GN   SSA2 OR YLL024C OR L0931.
OS   SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
OC   EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOMYCETES.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288C;
RX   MEDLINE; 89128457.
RA   SLATER M.R., CRAIG E.A.;
RL   NUCLEIC ACIDS RES. 17:805-806(1989).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288C;
RA   PURNELLE B., GOFFEAU A.;
RL   SUBMITTED (APR-1996) TO EMBL/GENBANK/DDBJ DATA BANKS.
RN   [3]
RP   SEQUENCE OF 71-638 FROM N.A.
RA   DUESTERHOEFT A., FLOETH M., HEUSS-NEITZEL D., HILBERT H., MOESTL D.;
RL   SUBMITTED (MAY-1996) TO EMBL/GENBANK/DDBJ DATA BANKS.
RN   [4]
RP   SEQUENCE OF 91-97 AND 325-341.
RC   STRAIN=S288C;
RX   MEDLINE; 95203288.
RA   GARRELS J.I., FUTCHER B., KOBAYASHI R., LATTER G.I., SCHWENDER B.,
RA   VOLPE T., WARNER J.R., MCLAUGHLIN C.S.;
RL   ELECTROPHORESIS 15:1466-1486(1994).
RN   [5]
RP   SEQUENCE OF 186-195.
RC   STRAIN=ATCC 38531 / Y41;
RA   NORBECK J., BLOMBERG A.;
RL   FEMS MICROBIOL. LETT. 137:1-8(1996).
RN   [6]
RP   ACETYLATION, AND PHOSPHORYLATION.
RA   GARRELS J.I., FUTCHER B., KOBAYASHI R., LATTER G.I., SCHWENDER B.,
RA   VOLPE T., WARNER J.R., MCLAUGHLIN C.S.;
RL   SUBMITTED (SEP-1994) TO THE SWISS-PROT DATA BANK.
CC   -!- FUNCTION: SSA2 MAY PLAY A ROLE IN THE TRANSPORT OF POLYPEPTIDES
CC       BOTH ACROSS THE MITOCHONDRIAL MEMBRANES AND INTO THE
CC       ENDOPLASMATIC RETICULUM. A FUNCTIONAL DIFFERENCE BETWEEN YEAST
CC       SSA1 AND SSA2 PROTEINS IS EXPECTED. SSA2 CAN PARTICIPATE IN THE
CC       ATP-DEPENDENT DISASSEMBLY OF CLATHRIN-COATED VESICLES.
CC   -!- PTM: PHOSPHORYLATED.
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC   -!- SIMILARITY: BELONGS TO HEAT SHOCK PROTEIN 70 FAMILY.
DR   EMBL; X12927; G4546; -.
DR   EMBL; Z73129; E245749; -.
DR   EMBL; X97560; E238712; -.
DR   PIR; S20139; S20139.
DR   HSSP; P19120; 1ATR.
DR   SWISS-2DPAGE; P10592; YEAST.
DR   YEPD; 9800; -.
DR   LISTA; SC01203; SSA2.
DR   SGD; L0002070; SSA2.
DR   PROSITE; PS00297; HSP70_1.
DR   PROSITE; PS00329; HSP70_2.
DR   PROSITE; PS01036; HSP70_3.
KW   HEAT SHOCK; ATP-BINDING; MULTIGENE FAMILY; ACETYLATION;
KW   PHOSPHORYLATION.
FT   INIT_MET      0      0
FT   MOD_RES       1      1       ACETYLATION.
SQ   SEQUENCE   638 AA;  69338 MW;  676F4EE3 CRC32;

  Query Match           13.1%;  Score 84;  DB 5;  Length 638;
  Best Local Similarity  26.3%;  Pred. No. 6.89e-01;
  Matches    10;  Conservative   12;  Mismatches  15;  Indels   1;  Gaps   1;

Db    260 rtacerakrtlsssaqtsveidslfegidfytsitrar 297
          |  | ::    :|  | : :::: | :  |:|:|:| |
Qy     11 RQECGKPALPSASEEQVAQDMEG-FSAATFFTTISRNR 47


RESULT    7
ID   HS7C_DICDI     STANDARD;      PRT;   640 AA.
AC   P36415;
DT   01-JUN-1994 (REL. 29, CREATED)
DT   01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)
DT   01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE)
DE   HEAT SHOCK COGNATE PROTEIN (AGINACTIN).
GN   HSPB OR HSC70.
OS   DICTYOSTELIUM DISCOIDEUM (SLIME MOLD).
OC   EUKARYOTA; PROTOZOA; SARCOMASTIGOPHORA; SARCODINA; RHIZOPODA;
OC   EUMYCETOZOA; DICTYOSTELIA.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AX3;
RX   MEDLINE; 94008983.
RA   HAUS U., TROMMLER P., FISHER P.R., HARTMANN H., LOTTSPEICH F.,
RA   NOEGEL A.A., SCHLEICHER M.;
```

```
RL   EMBO J. 12:3763-3771(1993).
RN   [2]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 34-46; 102-106 AND 298-309.
RC   STRAIN=AX3;
RX   MEDLINE; 94043116.
RA   EDDY R.J., SAUTERER R.A., CONDEELIS J.S.;
RL   J. BIOL. CHEM. 268:23267-23274(1993).
CC   -!- FUNCTION: AFFECTS ACTIN POLYMERIZATION THROUGH INTERACTION WITH
CC       THE ACTIN-BINDING PROTEIN CAP32/34. ACTS AS A CHAPERONE BY
CC       STIMULATING THE REFOLDING OF DENATURATED CAP32 AND CAP34.
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC. FOUND IN F-ACTIN-RICH REGIONS
CC       OF THE CELL CORTEX AND CELL PROTRUSIONS.
CC   -!- DEVELOPMENTAL STAGE: HEAT SHOCK COGNATE PROTEINS ARE EXPRESSED
CC       CONSTITUTIVELY DURING NORMAL DEVELOPMENT.
CC   -!- SIMILARITY: BELONGS TO HEAT SHOCK PROTEIN 70 FAMILY.
DR   EMBL; S65739; G415582; -.
DR   EMBL; X75263; G433875; -.
DR   EMBL; L22736; G433180; -.
DR   PIR; S37394; S37394.
DR   HSSP; P19120; 1NGA.
DR   DICTYDB; DD0????; HSPB.
DR   PROSITE; PS00297; HSP70_1.
DR   PROSITE; PS00329; HSP70_2.
DR   PROSITE; PS01036; HSP70_3.
KW   ATP-BINDING; CHAPERONE.
FT   CONFLICT      1     29       MSSIGIDLGTTYSCVGVWQNDRVEIIAND ->
FT                                IHHHINGNATWVVESGPVSEVLSFN (IN REF. 2).
FT   CONFLICT     32     32       N -> T (IN REF. 2).
FT   CONFLICT     64     64       V -> A (IN REF. 2).
FT   CONFLICT    180    180       R -> A (IN REF. 2).
FT   CONFLICT    237    237       S -> A (IN REF. 2).
FT   CONFLICT    240    240       V -> A (IN REF. 2).
FT   CONFLICT    341    341       I -> L (IN REF. 2).
FT   CONFLICT    352    352       F -> P (IN REF. 2).
SQ   SEQUENCE   640 AA;  70499 MW;  8550670E CRC32;

  Query Match           13.1%;  Score 84;  DB 5;  Length 640;
  Best Local Similarity  26.3%;  Pred. No. 6.89e-01;
  Matches    10;  Conservative   12;  Mismatches  15;  Indels   1;  Gaps   1;

Db    262 rtacerakrtlsssaqasieidslfegidfytsitrar 299
          |  | ::    :|  | : :::: | :  |:|:|:| |
Qy     11 RQECGKPALPSASEEQVAQDMEG-FSAATFFTTISRNR 47


RESULT    8
ID   HS71_YEAST     STANDARD;      PRT;   641 AA.
AC   P10591;
DT   01-JUL-1989 (REL. 11, CREATED)
DT   01-OCT-1994 (REL. 30, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   HEAT SHOCK PROTEIN SSA1 (HEAT SHOCK PROTEIN YG100).
GN   SSA1 OR YAL005C.
OS   SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
OC   EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOMYCETES.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288C;
RX   MEDLINE; 89128457.
RA   SLATER M.R., CRAIG E.A.;
RL   NUCLEIC ACIDS RES. 17:805-806(1989).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288C / AB972;
RX   MEDLINE; 95028152.
RA   CLARK M.W., KENG T., STORMS R.K., ZHONG W., FORTIN N., ZENG B.,
RA   DELANEY S., OUELLETTE B.F.F., BARTON A.B., KABACK D.B., BUSSEY H.;
RL   YEAST 10:535-541(1994).
RN   [3]
RP   SEQUENCE OF 590-641 FROM N.A.
RX   MEDLINE; 85087943.
RA   OGDEN R.C., LEE M.-C., KNAPP G.;
RL   NUCLEIC ACIDS RES. 12:9367-9382(1984).
RN   [4]
RP   REVISIONS TO 207; 417 AND 421.
RA   SLATER M.R.;
RL   SUBMITTED (JUN-1993) TO EMBL/GENBANK/DDBJ DATA BANKS.
RN   [5]
RP   SEQUENCE OF 91-97 AND 325-341.
RC   STRAIN=S288C;
RX   MEDLINE; 95203288.
RA   GARRELS J.I., FUTCHER B., KOBAYASHI R., LATTER G.I., SCHWENDER B.,
RA   VOLPE T., WARNER J.R., MCLAUGHLIN C.S.;
RL   ELECTROPHORESIS 15:1466-1486(1994).
RN   [6]
RP   SEQUENCE OF 186-195.
RC   STRAIN=ATCC 38531 / Y41;
RA   NORBECK J., BLOMBERG A.;
RL   FEMS MICROBIOL. LETT. 137:1-8(1996).
RN   [7]
RP   ACETYLATION.
RA   GARRELS J.I., FUTCHER B., KOBAYASHI R., LATTER G.I., SCHWENDER B.,
RA   VOLPE T., WARNER J.R., MCLAUGHLIN C.S.;
RL   SUBMITTED (SEP-1994) TO THE SWISS-PROT DATA BANK.
CC   -!- FUNCTION: SSA1 MAY PLAY A ROLE IN THE TRANSPORT OF POLYPEPTIDES
CC       BOTH ACROSS THE MITOCHONDRIAL MEMBRANES AND INTO THE
CC       ENDOPLASMATIC RETICULUM. A FUNCTIONAL DIFFERENCE BETWEEN YEAST
CC       SSA1 AND SSA2 PROTEINS IS EXPECTED. SSA1 CAN PARTICIPATE IN THE
CC       ATP-DEPENDENT DISASSEMBLY OF CLATHRIN-COATED VESICLES.
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC   -!- SIMILARITY: BELONGS TO HEAT SHOCK PROTEIN 70 FAMILY.
DR   EMBL; X12926; G312352; -.
DR   EMBL; L22015; G349747; ALT_SEQ.
DR   PIR; S25438; HHBYA1.
DR   PIR; S42164; S42164.
DR   HSSP; P19120; 1NGJ.
DR   SWISS-2DPAGE; P10591; YEAST.
DR   YEPD; 9788; -.
DR   LISTA; SC01202; SSA1.
DR   SGD; L0002069; SSA1.
DR   PROSITE; PS00297; HSP70_1.
DR   PROSITE; PS00329; HSP70_2.
DR   PROSITE; PS01036; HSP70_3.
KW   HEAT SHOCK; ATP-BINDING; MULTIGENE FAMILY; ACETYLATION.
FT   INIT_MET      0      0
FT   MOD_RES       1      1       ACETYLATION.
SQ   SEQUENCE   641 AA;  69526 MW;  5B23162E CRC32;

  Query Match           13.1%;  Score 84;  DB 5;  Length 641;
  Best Local Similarity  26.3%;  Pred. No. 6.89e-01;
  Matches    10;  Conservative   12;  Mismatches  15;  Indels   1;  Gaps   1;

Db    260 rtacerakrtlsssaqtsveidslfegidfytsitrar 297
          |  | ::    :|  | : :::: | :  |:|:|:| |
Qy     11 RQECGKPALPSASEEQVAQDMEG-FSAATFFTTISRNR 47


RESULT    9
ID   HS76_PIG       STANDARD;      PRT;   643 AA.
AC   Q04967;
DT   01-FEB-1994 (REL. 28, CREATED)
DT   01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)
DT   01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)
DE   HEAT SHOCK 70 KD PROTEIN 6 (HEAT SHOCK 70 KD PROTEIN B').
GN   HSP70B'.
OS   SUS SCROFA (PIG).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; ARTIODACTYLA.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LYMPHOCYTES;
RX   MEDLINE; 93326632.
RA   DEZEURE F., VAIMAN M., CHARDON P.;
RL   BIOCHIM. BIOPHYS. ACTA 1174:17-26(1993).
CC   -!- INDUCTION: BY HEAT SHOCK.
CC   -!- SIMILARITY: BELONGS TO HEAT SHOCK PROTEIN 70 FAMILY.
```

```
DR   EMBL; X68213; G1978; -.
DR   PIR; S25585; S25585.
DR   PIR; S34625; S34625.
DR   HSSP; P19120; 1NGI.
DR   PROSITE; PS00297; HSP70_1.
DR   PROSITE; PS00329; HSP70_2.
DR   PROSITE; PS01036; HSP70_3.
KW   ATP-BINDING; HEAT SHOCK; MULTIGENE FAMILY.
SQ   SEQUENCE   643 AA;  71109 MW;  6F2B78FD CRC32;

  Query Match           13.1%;  Score 84;  DB 5;  Length 643;
  Best Local Similarity  26.3%;  Pred. No. 6.89e-01;
  Matches    10;  Conservative   12;  Mismatches  15;  Indels   1;  Gaps   1;

Db    266 rtacerakrtlssstqatleidslfegvdfytsitrar 303
          |  | ::    :|  | : :::: | :  |:|:|:| |
Qy     11 RQECGKPALPSASEEQVAQDMEG-FSAATFFTTISRNR 47


RESULT   10
ID   HS70_XENLA     STANDARD;      PRT;   647 AA.
AC   P02827;
DT   21-JUL-1986 (REL. 01, CREATED)
DT   21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)
DT   01-JUL-1993 (REL. 26, LAST ANNOTATION UPDATE)
DE   HEAT SHOCK 70 KD PROTEIN (HSP70).
OS   XENOPUS LAEVIS (AFRICAN CLAWED FROG).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; AMPHIBIA; ANURA.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 85076567.
RA   BIENZ M.;
RL   EMBO J. 3:2477-2483(1984).
RN   [2]
RP   SEQUENCE OF 81-120 FROM N.A.
RX   MEDLINE; 84221917.
RA   BIENZ M.;
RL   PROC. NATL. ACAD. SCI. U.S.A. 81:3138-3142(1984).
CC   -!- SIMILARITY: BELONGS TO HEAT SHOCK PROTEIN 70 FAMILY.
DR   EMBL; X01102; G64796; -.
DR   EMBL; K02307; G214272; -.
DR   PIR; A03310; HHXL70.
DR   HSSP; P19120; 1ATR.
DR   PROSITE; PS00297; HSP70_1.
DR   PROSITE; PS00329; HSP70_2.
DR   PROSITE; PS01036; HSP70_3.
KW   ATP-BINDING; HEAT SHOCK.
SQ   SEQUENCE   647 AA;  70915 MW;  F8D46C32 CRC32;

  Query Match           13.1%;  Score 84;  DB 5;  Length 647;
  Best Local Similarity  26.3%;  Pred. No. 6.89e-01;
  Matches    10;  Conservative   12;  Mismatches  15;  Indels   1;  Gaps   1;

Db    265 rtacdrakrtlssssqasieidslfegidfytaitrar 302
          |  | ::    :|  | : :::: | :  |:|:|:| |
Qy     11 RQECGKPALPSASEEQVAQDMEG-FSAATFFTTISRNR 47


RESULT   11
ID   HS70_BLAEM     STANDARD;      PRT;   649 AA.
AC   P48720;
DT   01-FEB-1996 (REL. 33, CREATED)
DT   01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
DT   01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)
DE   HEAT SHOCK 70 KD PROTEIN.
GN   HSP70.
OS   BLASTOCLADIELLA EMERSONII.
OC   EUKARYOTA; FUNGI; MASTIGOMYCOTINA; CHYTRIDIOMYCETES.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 95129910.
RA   STEFANI R.M., GOMES S.L.;
RL   GENE 152:19-26(1995).
CC   -!- SIMILARITY: BELONGS TO HEAT SHOCK PROTEIN 70 FAMILY.
DR   EMBL; L22497; G773670; -.
KW   ATP-BINDING; CHAPERONE.
SQ   SEQUENCE   649 AA;  70833 MW;  8ACFF56F CRC32;

  Query Match           13.1%;  Score 84;  DB 5;  Length 649;
  Best Local Similarity  26.3%;  Pred. No. 6.89e-01;
  Matches    10;  Conservative   12;  Mismatches  15;  Indels   1;  Gaps   1;

Db    267 rtacerakrtlsssaqtsieidslfegidfytsitrar 304
          |  | ::    :|  | : :::: | :  |:|:|:| |
Qy     11 RQECGKPALPSASEEQVAQDMEG-FSAATFFTTISRNR 47


RESULT   12
ID   HS70_DAUCA     STANDARD;      PRT;   655 AA.
AC   P26791;
DT   01-AUG-1992 (REL. 23, CREATED)
DT   01-AUG-1992 (REL. 23, LAST SEQUENCE UPDATE)
DT   01-AUG-1992 (REL. 23, LAST ANNOTATION UPDATE)
DE   HEAT SHOCK 70 KD PROTEIN.
GN   HSP70.
OS   DAUCUS CAROTA (CARROT).
OC   EUKARYOTA; PLANTA; EMBRYOPHYTA; ANGIOSPERMAE; DICOTYLEDONEAE; APIALES;
OC   UMBELLIFERAE.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 92032789.
RA   LIN X., CHERN M., ZIMMERMAN J.L.;
RL   PLANT MOL. BIOL. 17:1245-1249(1991).
CC   -!- SIMILARITY: BELONGS TO HEAT SHOCK PROTEIN 70 FAMILY.
DR   EMBL; X60088; G18357; -.
DR   PIR; S18349; S18349.
DR   HSSP; P19120; 1NGJ.
DR   PROSITE; PS00297; HSP70_1.
DR   PROSITE; PS00329; HSP70_2.
DR   PROSITE; PS01036; HSP70_3.
KW   ATP-BINDING; HEAT SHOCK.
SQ   SEQUENCE   655 AA;  72051 MW;  791B240F CRC32;

  Query Match           13.1%;  Score 84;  DB 5;  Length 655;
  Best Local Similarity  30.6%;  Pred. No. 6.89e-01;
  Matches    11;  Conservative   14;  Mismatches   9;  Indels   2;  Gaps   2;

Db    272 erakrtlsssta-qttieidslyegvdfyttitrar 306
          | :| :|:|::  | : :::: : :  |:|||:| |
Qy     13 ECGKPALPSASEEQVAQDMEG-FSAATFFTTISRNR 47


RESULT   13
ID   BAT2_HUMAN     STANDARD;      PRT;  2142 AA.
AC   P48634;
DT   01-FEB-1996 (REL. 33, CREATED)
DT   01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
DT   01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)
DE   LARGE PROLINE-RICH PROTEIN BAT2 (HLA-B-ASSOCIATED TRANSCRIPT 2).
GN   BAT2.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=T-CELL;
RX   MEDLINE; 90192810.
RA   BANERJI J., SANDS J., STROMINGER J.L., SPIES T.;
RL   PROC. NATL. ACAD. SCI. U.S.A. 87:2374-2378(1990).
RN   [2]
RP   SEQUENCE OF 1-1860 FROM N.A.
RX   MEDLINE; 93272029.
RA   IRIS F.J.M., BOUGUELERET L., PRIEUR S., CATERINA D., PRIMAS G.,
RA   PERROT V., JURKA J., RODRIGUEZ-TOME P., CLAVERIE J.-M., DAUSSET J.,
RA   COHEN D.;
RL   NAT. GENET. 3:137-145(1993).
```

```
CC   -!- FUNCTION: UNKNOWN.
CC   -!- TISSUE SPECIFICITY: LIMITED TO CELL-LINES OF LEUKAEMIC ORIGIN.
DR   EMBL; M33509; G179339; -.
DR   EMBL; M33518; G179345; -.
DR   EMBL; M33512; G179345; JOINED.
DR   EMBL; Z15025; G29375; -.
DR   PIR; B35098; B35098.
DR   PIR; S36152; S36152.
DR   MIM; 142580; -.
KW   REPEAT.
FT   DOMAIN      519    524       POLY-PRO.
FT   DOMAIN      636    657       GLN-RICH.
FT   DOMAIN      684    688       POLY-PRO.
FT   DOMAIN      699    704       POLY-PRO.
FT   DOMAIN      814    821       POLY-PRO.
FT   DOMAIN     1340   1345       POLY-GLY.
FT   DOMAIN     1398   1403       POLY-GLY.
FT   DOMAIN     1436   1442       POLY-PRO.
FT   DOMAIN     1982   1991       POLY-PRO.
FT   DOMAIN       41   1795       4 X 57 AA TYPE A REPEATS.
FT   REPEAT       41     95       1-1.
FT   REPEAT       98    154       1-2.
FT   REPEAT      281    337       1-3.
FT   REPEAT     1740   1795       1-4.
FT   DOMAIN      337    549       2 X TYPE B REPEATS.
FT   REPEAT      337    418       2-1.
FT   REPEAT      476    549       2-2.
FT   DOMAIN     1899   2089       3 X 50 AA TYPE C REPEATS.
FT   REPEAT     1899   1948       3-1.
FT   REPEAT     1965   2014       3-2.
FT   REPEAT     2040   2089       3-3.
FT   CONFLICT     57     57       R -> A (IN REF. 2).
FT   CONFLICT    109    109       Q -> S (IN REF. 2).
FT   CONFLICT    414    414       P -> PPHRGPAGNWGPP (IN REF. 2).
FT   CONFLICT    532    532       T -> K (IN REF. 2).
FT   CONFLICT    682    682       Q -> K (IN REF. 2).
FT   CONFLICT    730    730       E -> D (IN REF. 2).
FT   CONFLICT    750    750       L -> R (IN REF. 2).
FT   CONFLICT    834    834       A -> T (IN REF. 2).
FT   CONFLICT   1035   1035       G -> A (IN REF. 2).
FT   CONFLICT   1068   1068       M -> L (IN REF. 2).
FT   CONFLICT   1285   1285       P -> R (IN REF. 2).
FT   CONFLICT   1400   1400       G -> A (IN REF. 2).
FT   CONFLICT   1611   1611       T -> S (IN REF. 2).
FT   CONFLICT   1729   1729       G -> A (IN REF. 2).
SQ   SEQUENCE   2142 AA;  227840 MW;  2CFEF88A CRC32;

  Query Match           13.1%;  Score 84;  DB 1;  Length 2142;
  Best Local Similarity  27.8%;  Pred. No. 6.89e-01;
  Matches    15;  Conservative   13;  Mismatches  26;  Indels   0;  Gaps   0;

Db   1281 apappraaakspdlsnqnsdqaneewetasessdftserrgdkeapppvlltpk 1334
          :|:|||     | |:: : :|  :: |  | :: ||:  |  :    | | | :
Qy      6 GPGPPRQECGKPALPSASEEQVAQDMEGFSAATFFTTISRNRRLKGRPPLPTQR 59


RESULT   14
ID   GOLI_DROME     STANDARD;      PRT;   284 AA.
AC   Q06003;
DT   01-FEB-1994 (REL. 28, CREATED)
DT   01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)
DT   01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)
DE   GOLIATH PROTEIN (G1 PROTEIN).
GN   GOL OR G1 OR GL.
OS   DROSOPHILA MELANOGASTER (FRUIT FLY).
OC   EUKARYOTA; METAZOA; ARTHROPODA; INSECTA; DIPTERA.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=EMBRYO;
RX   MEDLINE; 93216124.
RA   BOUCHARD M.L., COTE S.;
RL   GENE 125:205-209(1993).
CC   -!- FUNCTION: REGULATION OF GENE EXPRESSION DURING MESODERM FORMATION.
CC       PUTATIVE ROLE AS TRANSCRIPTION FACTOR.
CC   -!- TISSUE SPECIFICITY: VISCERAL MESODERM AND PRIMORDIA OF SOMATIC
CC       MUSCULATURE.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- SIMILARITY: CONTAINS A C3HC4-CLASS ZINC FINGER (C3H2C3 SUBFAMILY).
DR   EMBL; M97204; G157535; -.
DR   PIR; JC1495; JC1495.
DR   FLYBASE; FBGN0004919; GOL.
KW   DEVELOPMENTAL PROTEIN; ZINC-FINGER; TRANSCRIPTION REGULATION;
KW   DNA-BINDING; NUCLEAR PROTEIN.
FT   ZN_FING     126    166       C3HC4-TYPE.
FT   DOMAIN      208    259       GLN/PRO/SER-RICH.
SQ   SEQUENCE   284 AA;  31973 MW;  FB98BE96 CRC32;

  Query Match           12.9%;  Score 83;  DB 4;  Length 284;
  Best Local Similarity  26.3%;  Pred. No. 9.52e-01;
  Matches    15;  Conservative   16;  Mismatches  24;  Indels   2;  Gaps   2;

Db    219 asnmssfapshyfqs-srspsssvqqqlapltyqphpqqaasergrrn-sapatmph 273
          | :| :|::: :| : ||:     :  |:     | |   |: :||:: |:|:   |
Qy     28 AQDMEGFSAATFFTTISRNRRLKGRPPLPTQRWSPCPSNLAAPWGRWDGSSPSPGRH 84


RESULT   15
ID   FUCO_RAT       STANDARD;      PRT;   462 AA.
AC   P17164;
DT   01-AUG-1990 (REL. 15, CREATED)
DT   01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)
DT   01-JUN-1994 (REL. 29, LAST ANNOTATION UPDATE)
DE   ALPHA-L-FUCOSIDASE PRECURSOR (EC 3.2.1.51) (ALPHA-L-FUCOSIDE
DE   FUCOHYDROLASE).
GN   FUCA1.
OS   RATTUS NORVEGICUS (RAT).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; RODENTIA.
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 29-40; 90-124 AND 307-372.
RC   TISSUE=LIVER;
RX   MEDLINE; 90147522.
RA   FISHER K.J., ARONSON N.N. JR.;
RL   BIOCHEM. J. 264:695-701(1989).
CC   -!- FUNCTION: ALPHA-L-FUCOSIDASE IS RESPONSIBLE FOR HYDROLYSING THE
CC       ALPHA-1,6-LINKED FUCOSE JOINED TO THE REDUCING-END
CC       N-ACETYLGLUCOSAMINE OF THE CARBOHYDRATE MOIETIES OF GLYCOPROTEINS.
CC   -!- CATALYTIC ACTIVITY: AN ALPHA-L-FUCOSIDE + H(2)O = AN ALCOHOL +
CC       L-FUCOSE.
CC   -!- SUBUNIT: HOMOTETRAMER.
CC   -!- SUBCELLULAR LOCATION: LYSOSOMAL.
CC   -!- SIMILARITY: BELONGS TO FAMILY 29 OF GLYCOSYL HYDROLASES.
DR   EMBL; X16145; G55651; -.
DR   PIR; S07074; S07074.
DR   PIR; S10235; S10235.
DR   PROSITE; PS00385; ALPHA_L_FUCOSIDASE.
KW   HYDROLASE; GLYCOSIDASE; LYSOSOME; GLYCOPROTEIN; SIGNAL.
FT   SIGNAL        1     28
FT   CHAIN        29    462       ALPHA-L-FUCOSIDASE.
FT   ACT_SITE    292    292       MAY BE IMPORTANT FOR CATALYSIS.
FT   CARBOHYD    237    237       POTENTIAL.
FT   CARBOHYD    264    264       POTENTIAL.
FT   CARBOHYD    378    378       POTENTIAL.
SQ   SEQUENCE   462 AA;  53486 MW;  4923CBE4 CRC32;

  Query Match           12.9%;  Score 83;  DB 4;  Length 462;
  Best Local Similarity  46.2%;  Pred. No. 9.52e-01;
  Matches    12;  Conservative    5;  Mismatches   7;  Indels   2;  Gaps   2;

Db    280 rphslpdhkwemctsvdkaswgyrrd 305
          || :|| ::|  | |   |:|| |:|
Qy     52 RP-PLPTQRWSPCPSNLAAPWG-RWD 75


Search completed: Wed Aug 20 11:10:10 1997
Job time : 16 secs.
```



Release 2.1D John F. Collins, Biocomputing Research Unit.

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Wed Aug 20 11:12:02 1997; MasPar time 5.48 Seconds
438.125 Million cell updates/sec
Tabular output not generated.

Title: >US-08-320-157-22
Description: (1-210) from US08320157.pep
Perfect Score: 1554
Sequence: 1 MASGQGPGPPRQECGEPALP.....LVVLGVVLLGQFVVRRFFKS 210

Scoring table: PAM 150
Gap 11

Searched: 96640 seqs, 11439865 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: a-geneseq27
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20

Statistics: Mean 32.457; Variance 141.622; scale 0.229

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description	Pred. No.
1	1543	99.3	211	20	W03668	Bak protein.	4.01e-136
2	1543	99.3	211	17	R81451	Bcl-Y apoptosis-relat	4.01e-136
3	1543	99.3	211	13	R77876	Human Cdn-1.	4.01e-136
4	1511	97.2	211	20	W03669	Bak-2 protein.	5.36e-133
5	1511	97.2	211	13	R77877	Human Cdn-2.	5.36e-133
6	1136	73.1	152	13	R77879	Human Cdn-1(60-211).	1.82e-96
7	1061	68.3	141	13	R77880	Human Cdn-1(71-211).	3.42e-89
8	880	56.6	116	13	R77881	Human Cdn-1(96-211).	1.07e-71
9	258	16.6	190	13	R68884	Chicken lymphoid BCL-	3.49e-13
10	256	16.5	232	19	W01020	Apoptosis-blocking pr	5.25e-13
11	250	16.1	233	13	R68887	Human thymus BCL-XL.	1.79e-12
12	250	16.1	233	20	W05821	Bcl-XL protein.	1.79e-12
13	245	15.8	232	19	W01019	Apoptosis-blocking pr	4.95e-12
14	245	15.8	239	19	W01018	Apoptosis-blocking pr	4.95e-12
15	243	15.6	205	13	R71405	Human bcl-2 beta prot	7.43e-12
16	243	15.6	205	13	R68886	Human thymus BCL-2.	7.43e-12
17	243	15.6	205	13	R70332	Human bcl-2 protein.	7.43e-12
18	243	15.6	239	13	R71404	Human bcl-2 alpha pro	7.43e-12
19	243	15.6	239	9	R47344	Human oncogene bcl-2	7.43e-12
20	243	15.6	239	1	P80987	Sequence of bcl-2-alp	7.43e-12
21	243	15.6	239	13	R70331	Human bcl-2 protein.	7.43e-12
22	243	15.6	239	8	R42312	Bcl-2 oncogene produc	7.43e-12
23	213	13.7	88	13	R77878	Human Cdn-3.	3.14e-09
24	213	13.7	205	1	P80988	Sequence of bcl-2-bet	3.14e-09
25	204	13.1	192	13	R71407	Murine Bax protein.	1.88e-08
26	203	13.1	229	19	W01021	Apoptosis-blocking pr	2.29e-08
27	194	12.5	192	13	R71406	Human Bax protein.	1.36e-07
28	179	11.5	350	12	R68814	Human mcl-1 gene prod	2.55e-06
29	150	9.7	63	13	R68885	Human thymus BCL-X1.	6.50e-04
30	98	6.3	277	14	R76996	Deduced sequence enco	7.07e+00
31	97	6.2	170	13	R68888	Human thymus BCL-XS.	8.35e+00
32	96	6.2	443	19	W05298	Esterase secretory pr	9.86e+00
33	94	6.0	338	14	R73955	Duffy blood group gpD	1.37e+01
34	91	5.9	2749	3	R13887	Inositol-3-phosphate	2.24e+01
35	89	5.7	277	14	R74737	ACT-4 cell surface re	3.10e+01
36	89	5.7	277	15	R79904	ACT-4-h-1 receptor se	3.10e+01
37	89	5.7	471	14	R79163	Partial sequence of b	3.10e+01
38	88	5.7	912	14	R82658	Human mGluR4.	3.64e+01
39	88	5.7	912	13	R72092	Human mGluR4.	3.64e+01
40	88	5.7	2695	20	W07632	Human type I inositol	3.64e+01
41	88	5.7	2713	19	W00168	Human T cell inositol	3.64e+01
42	87	5.6	451	14	R79909	N. crassa mtr gene pr	4.27e+01
43	87	5.6	470	9	R47118	Mtr protein of Neuros	4.27e+01
44	87	5.6	1984	18	R99639	Peripheral nervous sy	4.27e+01
45	87	5.6	1989	18	R92317	Peripheral nervous sy	4.27e+01

ALIGNMENTS

RESULT 1
ID W03668 standard; Protein; 211 AA.
AC W03668;
DT 22-FEB-1997 (first entry)
DE Bak protein.
KW Human; Bak; apoptosis; latency; virus replication;
KW Epstein-Barr virus; BHRF1; fusion protein; epitope tag;
KW drug screening; co-precipitation; ELISA; immunoassay; antibody;
KW protein interactive trapping; virucide; antitumour; diagnostic.
OS Homo sapiens.
PN WO9633416-A1.
PD 24-OCT-1996.
PF 19-APR-1996; U05639.
PR 20-APR-1995; US-426529.
PA (LXRB-) LXR BIOTECHNOLOGY INC.
PI Barr PJ, Kiefer MC;
DR WPI; 96-485886/48.
DR N-PSDB; T42138.
PT Screening for anti-viral agents - by detecting the ability of an
PT agent to disrupt the interaction of a Bak protein and a viral
PT protein
PS Disclosure; Fig 1; 24pp; English.
CC This Bak protein sequence represents a bcl-1 homologue which
CC interacts with Epstein-Barr virus (EBV) early lytic cycle BHRF1
CC protein, and is capable of modulating apoptosis. The protein may
CC be used in complete or partial form, or as an epitope tag fusion
CC protein, in a new virucide drug screening method, which involves
CC combination of Bak protein and a viral protein (e.g. EBV BHRF1),
CC exposure to a test compound, and monitoring for disruption of the
CC interaction, e.g. by co-precipitation, protein interactive trapping
CC or ELISA. Interaction of Bak and viral proteins allows viral
CC replication or latency in the absence of apoptosis. Compounds which
CC inhibit the interaction may be used as virucide, antitumour or
CC diagnostic agents.
SQ Sequence 211 AA;

Query Match 99.3%; Score 1543; DB 20; Length 211;
Best Local Similarity 99.5%; Pred. No. 4.01e-136;
Matches 210; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Db 1 masgqgpgpprqecgepalpsaseeqvaqdteevfrsyvfyrhqqeqeaegvaapadpem 60
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1 MASGQGPGPPRQECGEPALPSASEEQVAQDTEEVFRSYVFYRHQQEQEAEGVAAPADPEM 60

```
Db     61 vtlplqpsstmgqvgrqlaiigddinrrydsefqtmlqhlqptaenayeyftkiatslfe 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     61 VTLPLQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENAYEYFTKIATSLFE 120

Db    121 sginwgrvvallgfgyrlalhvyqhgltgflgqvtrfvvdfmlhhciarwiaqrggwvaa 180
          || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    121 SG-NWGRVVALLGFGYRLALHVYQHGLTGFLGQVTRFVVDFMLHHCIARWIAQRGGWVAA 179

Db    181 lnlgngpilnvlvvlgvvllgqfvvrrffks 211
          |||||||||||||||||||||||||||||||
Qy    180 LNLGNGPILNVLVVLGVVLLGQFVVRRFFKS 210


RESULT    2
ID    R81451 standard; Protein; 211 AA.
AC    R81451;
DT    02-JUL-1996  (first entry)
DE    Bcl-Y apoptosis-related protein.
KW    Bcl-Y; apoptosis; cell proliferation; cell death; diagnosis;
KW    therapy.
OS    Homo sapiens.
FH    Key             Location/Qualifiers
FT    Domain          188..205
FT    /label= C-terminal_domain
FT    /note= "putative membrane localisation sequence"
PN    WO9605232-A1.
PD    22-FEB-1996.
PF    09-AUG-1995; U10103.
PR    09-AUG-1994; US-287427.
PR    11-OCT-1994; US-321071.
PA    (IMMU-) IMMUNOGEN INC.
PI    Chittenden TD;
DR    WPI; 96-139648/14.
DR    N-PSDB; T17375.
PT    New isolated human Bcl-Y protein - used to develop prods. for
PT    treating disorders characterised by inappropriate cell proliferation
PT    or cell death
PS    Claim 3; Fig 4; 100pp; English.
CC    Bcl-Y protein (R81451) is a member of the Bcl-2 family and can
CC    induce apoptosis in cells and function as a negative regulator of
CC    Bcl-2 function.  Bcl-Y mRNA was detected in all human tumour cell
CC    lines examined and is also widely expressed in primary human
CC    tissues.  It can be obtd. by expression of a full-length cDNA
CC    clone (T17375) in pref. mammalian host cells.  Bcl-Y can be used to
CC    develop prods. for treating disorders associated with inappropriate
CC    cell proliferation or cell death, and to raise antibodies used for
CC    the diagnosis or monitoring of such disorders.
SQ    Sequence   211 AA;

  Query Match           99.3%;  Score 1543;  DB 17;  Length 211;
  Best Local Similarity  99.5%;  Pred. No. 4.01e-136;
  Matches   210;  Conservative    0;  Mismatches   0;  Indels   1;  Gaps   1;

Db      1 masgqgpgpprqecgepalpsaseeqvaqdteevfrsyvfyrhqqeqeaegvaapadpem 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      1 MASGQGPGPPRQECGEPALPSASEEQVAQDTEEVFRSYVFYRHQQEQEAEGVAAPADPEM 60

Db     61 vtlplqpsstmgqvgrqlaiigddinrrydsefqtmlqhlqptaenayeyftkiatslfe 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     61 VTLPLQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENAYEYFTKIATSLFE 120

Db    121 sginwgrvvallgfgyrlalhvyqhgltgflgqvtrfvvdfmlhhciarwiaqrggwvaa 180
          || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    121 SG-NWGRVVALLGFGYRLALHVYQHGLTGFLGQVTRFVVDFMLHHCIARWIAQRGGWVAA 179

Db    181 lnlgngpilnvlvvlgvvllgqfvvrrffks 211
          |||||||||||||||||||||||||||||||
Qy    180 LNLGNGPILNVLVVLGVVLLGQFVVRRFFKS 210


RESULT    3
ID    R77876 standard; Protein; 211 AA.
AC    R77876;
DT    21-NOV-1995  (first entry)
DE    Human Cdn-1.
KW    Cdn-1; apoptosis modulator; adoptive immunotherapy; therapy; HIV;
KW    autoimmune disease; reperfusion injury; hepatitis, osteoporosis;
KW    shock; lymphoma; eczema.
OS    Homo sapiens.
PN    WO9515084-A.
PD    08-JUN-1995.
PF    30-NOV-1994; U13930.
PR    30-NOV-1993; US-160067.
PR    07-OCT-1994; US-320157.
PA    (LXRB-) LXR BIOTECHNOLOGY INC.
PI    Barr PJ,  Kiefer MC;
DR    WPI; 95-215106/28.
DR    N-PSDB; Q95492.
PT    New nucleic acid sequences encoding Cdn apoptosis modulators  - and
PT    related vectors, transformed cells, proteins and antibodies, useful
PT    or diagnosis and treatment e.g. of HIV infection, reperfusion injury
PT    etc.
PS    Disclosure; Fig.3A-B; 66pp; English.
CC    Cdn-1 cDNA was isolated from a human heart cDNA library using a
CC    previously isolated clone as probe.  Recombinant Cdn-1 was produced
CC    in Sf9 and human colon adenocarcinoma HT29 cells.  Expression of
CC    Cdn-1 in WI-L2 lymphoblastoid cells resulted in increased cell
CC    survival in response to anti-Fas-mediated apoptosis.
SQ    Sequence   211 AA;

  Query Match           99.3%;  Score 1543;  DB 13;  Length 211;
  Best Local Similarity  99.5%;  Pred. No. 4.01e-136;
  Matches   210;  Conservative    0;  Mismatches   0;  Indels   1;  Gaps   1;

Db      1 masgqgpgpprqecgepalpsaseeqvaqdteevfrsyvfyrhqqeqeaegvaapadpem 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      1 MASGQGPGPPRQECGEPALPSASEEQVAQDTEEVFRSYVFYRHQQEQEAEGVAAPADPEM 60

Db     61 vtlplqpsstmgqvgrqlaiigddinrrydsefqtmlqhlqptaenayeyftkiatslfe 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     61 VTLPLQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENAYEYFTKIATSLFE 120

Db    121 sginwgrvvallgfgyrlalhvyqhgltgflgqvtrfvvdfmlhhciarwiaqrggwvaa 180
          || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    121 SG-NWGRVVALLGFGYRLALHVYQHGLTGFLGQVTRFVVDFMLHHCIARWIAQRGGWVAA 179

Db    181 lnlgngpilnvlvvlgvvllgqfvvrrffks 211
          |||||||||||||||||||||||||||||||
Qy    180 LNLGNGPILNVLVVLGVVLLGQFVVRRFFKS 210


RESULT    4
ID    W03669 standard; Protein; 211 AA.
AC    W03669;
DT    22-FEB-1997  (first entry)
DE    Bak-2 protein.
KW    Human; Bak-2; apoptosis; latency; virus replication;
KW    Epstein-Barr virus; BHRF1; fusion protein; epitope tag;
KW    drug screening; co-precipitation; ELISA; immunoassay; antibody;
KW    protein interactive trapping; virucide; antitumour; diagnostic.
OS    Homo sapiens.
PN    WO9633416-A1.
PD    24-OCT-1996.
PF    19-APR-1996; U05639.
PR    20-APR-1995; US-426529.
PA    (LXRB-) LXR BIOTECHNOLOGY INC.
PI    Barr PJ,  Kiefer MC;
DR    WPI; 96-485886/48.
DR    N-PSDB; T42139.
PT    Screening for anti-viral agents - by detecting the ability of an
PT    agent to disrupt the interaction of a Bak protein and a viral
PT    protein
PS    Disclosure; Fig 2; 24pp; English.
CC    This Bak-2 protein sequence represents a bcl-1 homologue which
CC    interacts with Epstein-Barr virus (EBV) early lytic cycle BHRF1
CC    protein, and is capable of modulating apoptosis.  The protein may
```

```
CC   be used in complete or partial form, or as an epitope tag fusion
CC   protein, in a new virucide drug screening method, which involves
CC   combination of Bak-2 protein and a viral protein (e.g. EBV BHRF1),
CC   exposure to a test compound, and monitoring for disruption of the
CC   interaction, e.g. by co-precipitation, protein interactive trapping
CC   or ELISA.  Interaction of Bak-2 and viral proteins allows viral
CC   replication or latency in the absence of apoptosis.  Compounds which
CC   inhibit the interaction may be used as virucide, antitumour or
CC   diagnostic agents.
SQ   Sequence   211 AA;

  Query Match           97.2%;  Score 1511;  DB 20;  Length 211;
  Best Local Similarity  96.7%;  Pred. No. 5.36e-133;
  Matches   204;  Conservative    5;  Mismatches   1;  Indels   1;  Gaps   1;

Db       1 masgqgpgpprqecgepalpsaseeqvaqdteevfrsyvfyhhqqeqeaegaaapadpem 60
           ||||||||||||||||||||||||||||||||||||||||||:|||||||| ||||||||
Qy       1 MASGQGPGPPRQECGEPALPSASEEQVAQDTEEVFRSYVFYRHQQEQEAEGVAAPADPEM 60

Db      61 vtlplqpsstmgqvgrqlaiigddinrrydsefqtmlqhlqptaenayeyftkiasslfe 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||
Qy      61 VTLPLQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENAYEYFTKIATSLFE 120

Db     121 sginwgrvvallgfsyrlalhiyqrgltgflgqvtrfvvdfmlhhciarwiaqrggwvaa 180
           || ||||||||||:||||||:||:||||||||||||||||||||||||||||||||||||
Qy     121 SG-NWGRVVALLGFGYRLALHVYQHGLTGFLGQVTRFVVDFMLHHCIARWIAQRGGWVAA 179

Db     181 lnlgngpilnvlvvlgvvllgqfvvrrffks 211
           |||||||||||||||||||||||||||||||
Qy     180 LNLGNGPILNVLVVLGVVLLGQFVVRRFFKS 210


RESULT    5
ID   R77877 standard; Protein; 211 AA.
AC   R77877;
DT   21-NOV-1995  (first entry)
DE   Human Cdn-2.
KW   Cdn-2; apoptosis modulator; adoptive immunotherapy; therapy; HIV;
KW   autoimmune disease; reperfusion injury; hepatitis, osteoporosis;
KW   shock; lymphoma; eczema.
OS   Homo sapiens.
PN   WO9515084-A.
PD   08-JUN-1995.
PF   30-NOV-1994; U13930.
PR   30-NOV-1993; US-160067.
PR   07-OCT-1994; US-320157.
PA   (LXRB-) LXR BIOTECHNOLOGY INC.
PI   Barr PJ,  Kiefer MC;
DR   WPI; 95-215106/28.
DR   N-PSDB; Q95493.
PT   New nucleic acid sequences encoding Cdn apoptosis modulators  - and
PT   related vectors, transformed cells, proteins and antibodies, useful
PT   or diagnosis and treatment e.g. of HIV infection, reperfusion injury
PT   etc.
PS   Disclosure; Fig.5D-E; 66pp; English.
CC   Cdn-2 cDNA was isold. from a human placental genomic library
CC   using a 950 bp fragment of Cdn-1 cDNA.  Expression of Cdn-2
CC   in mouse progenitor B-cell FL5.12 cells decreased IL-3-induced
CC   apoptosis.  The Cdn-2 protein displayed 97% amino acid identity
CC   with Cdn-1 (R77876).
SQ   Sequence   211 AA;

  Query Match           97.2%;  Score 1511;  DB 13;  Length 211;
  Best Local Similarity  96.7%;  Pred. No. 5.36e-133;
  Matches   204;  Conservative    5;  Mismatches   1;  Indels   1;  Gaps   1;

Db       1 masgqgpgpprqecgepalpsaseeqvaqdteevfrsyvfyhhqqeqeaegaaapadpem 60
           ||||||||||||||||||||||||||||||||||||||||||:|||||||| ||||||||
Qy       1 MASGQGPGPPRQECGEPALPSASEEQVAQDTEEVFRSYVFYRHQQEQEAEGVAAPADPEM 60

Db      61 vtlplqpsstmgqvgrqlaiigddinrrydsefqtmlqhlqptaenayeyftkiasslfe 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||
Qy      61 VTLPLQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENAYEYFTKIATSLFE 120

Db     121 sginwgrvvallgfsyrlalhiyqrgltgflgqvtrfvvdfmlhhciarwiaqrggwvaa 180
           || ||||||||||:||||||:||:||||||||||||||||||||||||||||||||||||
Qy     121 SG-NWGRVVALLGFGYRLALHVYQHGLTGFLGQVTRFVVDFMLHHCIARWIAQRGGWVAA 179

Db     181 lnlgngpilnvlvvlgvvllgqfvvrrffks 211
           |||||||||||||||||||||||||||||||
Qy     180 LNLGNGPILNVLVVLGVVLLGQFVVRRFFKS 210


RESULT    6
ID   R77879 standard; Protein; 152 AA.
AC   R77879;
DT   21-NOV-1995  (first entry)
DE   Human Cdn-1(60-211).
KW   Cdn-1; apoptosis modulator; adoptive immunotherapy; therapy; HIV;
KW   autoimmune disease; reperfusion injury; hepatitis, osteoporosis;
KW   shock; lymphoma; eczema.
OS   Homo sapiens.
PN   WO9515084-A.
PD   08-JUN-1995.
PF   30-NOV-1994; U13930.
PR   30-NOV-1993; US-160067.
PR   07-OCT-1994; US-320157.
PA   (LXRB-) LXR BIOTECHNOLOGY INC.
PI   Barr PJ,  Kiefer MC;
DR   WPI; 95-215106/28.
PT   New nucleic acid sequences encoding Cdn apoptosis modulators  - and
PT   related vectors, transformed cells, proteins and antibodies, useful
PT   or diagnosis and treatment e.g. of HIV infection, reperfusion injury
PT   etc.
PS   Disclosure; Fig.11; 66pp; English.
CC   Expression of Cdn-1 in WI-L2 lymphoblastoid cells resulted in
CC   increased cell survival in response to anti-Fas-mediated apoptosis.
CC   Deletion of the N-terminal 59 amino acids of Cdn-1 only slightly
CC   decreased this activity, suggesting that small, truncated Cdn-1
CC   molecules may be potent therapeutics.
SQ   Sequence   152 AA;

  Query Match           73.1%;  Score 1136;  DB 13;  Length 152;
  Best Local Similarity  99.3%;  Pred. No. 1.82e-96;
  Matches   151;  Conservative    0;  Mismatches   0;  Indels   1;  Gaps   1;

Db       1 mvtlplqpsstmgqvgrqlaiigddinrrydsefqtmlqhlqptaenayeyftkiatslf 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      60 MVTLPLQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENAYEYFTKIATSLF 119

Db      61 esginwgrvvallgfgyrlalhvyqhgltgflgqvtrfvvdfmlhhciarwiaqrggwva 120
           ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     120 ESG-NWGRVVALLGFGYRLALHVYQHGLTGFLGQVTRFVVDFMLHHCIARWIAQRGGWVA 178

Db     121 alnlgngpilnvlvvlgvvllgqfvvrrffks 152
           |||||||||||||||||||||||||||||||
Qy     179 ALNLGNGPILNVLVVLGVVLLGQFVVRRFFKS 210


RESULT    7
ID   R77880 standard; Protein; 141 AA.
AC   R77880;
DT   21-NOV-1995  (first entry)
DE   Human Cdn-1(71-211).
KW   Cdn-1; apoptosis modulator; adoptive immunotherapy; therapy; HIV;
KW   autoimmune disease; reperfusion injury; hepatitis, osteoporosis;
KW   shock; lymphoma; eczema.
OS   Homo sapiens.
PN   WO9515084-A.
PD   08-JUN-1995.
PF   30-NOV-1994; U13930.
PR   30-NOV-1993; US-160067.
PR   07-OCT-1994; US-320157.
PA   (LXRB-) LXR BIOTECHNOLOGY INC.
PI   Barr PJ,  Kiefer MC;
DR   WPI; 95-215106/28.
```

```
PT   New nucleic acid sequences encoding Cdn apoptosis modulators  - and
PT   related vectors, transformed cells, proteins and antibodies, useful
PT   or diagnosis and treatment e.g. of HIV infection, reperfusion injury
PT   etc.
PS   Disclosure; Fig.11; 66pp; English.
CC   Expression of Cdn-1 in WI-L2 lymphoblastoid cells resulted in
CC   increased cell survival in response to anti-Fas-mediated apoptosis.
CC   Deletion of the N-terminal 70 amino acids of Cdn-1 improved this
CC   activity, suggesting that small, truncated Cdn-1 molecules may be
CC   potent therapeutics.
SQ   Sequence   141 AA;

  Query Match           68.3%;  Score 1061;  DB 13;  Length 141;
  Best Local Similarity  99.3%;  Pred. No. 3.42e-89;
  Matches   140;  Conservative    0;  Mismatches   0;  Indels   1;  Gaps   1;

Db      1 mgqvgrqlaiigddinrrydsefqtmlqhlqptaenayeyftkiatslfesginwgrvva 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Qy     71 MGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENAYEYFTKIATSLFESG-NWGRVVA 129

Db     61 llgfgyrlalhvyqhgltgflgqvtrfvvdfmlhhciarwiaqrggwvaalnlgngpiln 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    130 LLGFGYRLALHVYQHGLTGFLGQVTRFVVDFMLHHCIARWIAQRGGWVAALNLGNGPILN 189

Db    121 vlvvlgvvllgqfvvrrffks 141
          |||||||||||||||||||||
Qy    190 VLVVLGVVLLGQFVVRRFFKS 210


RESULT    8
ID   R77881 standard; Protein; 116 AA.
AC   R77881;
DT   21-NOV-1995  (first entry)
DE   Human Cdn-1(96-211).
KW   Cdn-1; apoptosis modulator; adoptive immunotherapy; therapy; HIV;
KW   autoimmune disease; reperfusion injury; hepatitis, osteoporosis;
KW   shock; lymphoma; eczema.
OS   Homo sapiens.
PN   WO9515084-A.
PD   08-JUN-1995.
PF   30-NOV-1994; U13930.
PR   30-NOV-1993; US-160067.
PR   07-OCT-1994; US-320157.
PA   (LXRB-) LXR BIOTECHNOLOGY INC.
PI   Barr PJ,  Kiefer MC;
DR   WPI; 95-215106/28.
PT   New nucleic acid sequences encoding Cdn apoptosis modulators  - and
PT   related vectors, transformed cells, proteins and antibodies, useful
PT   or diagnosis and treatment e.g. of HIV infection, reperfusion injury
PT   etc.
PS   Disclosure; Fig.11; 66pp; English.
CC   Expression of Cdn-1 in WI-L2 lymphoblastoid cells resulted in
CC   increased cell survival in response to anti-Fas-mediated apoptosis.
CC   Truncated Cdn-1 derivatives given in R77879-81 were used to
CC   test the effects of deleting the N-terminal sequences of Cdn-1
CC   on this activity.
SQ   Sequence   116 AA;

  Query Match           56.6%;  Score 880;  DB 13;  Length 116;
  Best Local Similarity  99.1%;  Pred. No. 1.07e-71;
  Matches   115;  Conservative    0;  Mismatches   0;  Indels   1;  Gaps   1;

Db      1 mlqhlqptaenayeyftkiatslfesginwgrvvallgfgyrlalhvyqhgltgflgqvt 60
          ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Qy     96 MLQHLQPTAENAYEYFTKIATSLFESG-NWGRVVALLGFGYRLALHVYQHGLTGFLGQVT 154

Db     61 rfvvdfmlhhciarwiaqrggwvaalnlgngpilnvlvvlgvvllgqfvvrrffks 116
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    155 RFVVDFMLHHCIARWIAQRGGWVAALNLGNGPILNVLVVLGVVLLGQFVVRRFFKS 210


RESULT    9
ID   R68884 standard; Protein; 190 AA.
AC   R68884;
DT   10-AUG-1995  (first entry)
DE   Chicken lymphoid BCL-X.
KW   Chicken; bird; fowl; BCL-X; apoptosis; cell death; cancer;
KW   neurodegenerative disease; autoimmune disease; Parkinson's disease;
KW   amylotrophic lateral sclerosis; multiple sclerosis; oncogene.
OS   Gallus domesticus.
PN   WO9500642-A.
PD   05-JAN-1995.
PF   22-JUN-1994; U07089.
PR   22-JUN-1993; US-081448.
PA   (ARCH-) ARCH DEV CORP.
PA   (UNMI ) UNIV MICHIGAN.
PI   Boise LH,  Nunez G,  Thompson CB;
DR   WPI; 95-052079/07.
DR   N-PSDB; Q81696.
PT   New poly-nucleotide encoding new poly-peptide(s) that modify
PT   apoptosis - and related vectors, recombinant cells and
PT   antibodies, useful in assay and for control of cell death in e.g.
PT   neuronal cells, lymphocytes and cancers
PS   Claim 4; Page 87; 127pp; English.
CC   This protein may be expressed recombinantly, particularly with pcmV
CC   plasmids as vectors for expression in mammalian cell cultures.
CC   The protein has particular application in cancer cells (failure of
CC   programmed cell death (PCD)) or neurodegenerative and autoimmune diseases
CC   (premature PCD), e.g. Parkinson's disease, amylotrophic lateral
CC   sclerosis and multiple sclerosis.
SQ   Sequence   190 AA;

  Query Match           16.6%;  Score 258;  DB 13;  Length 190;
  Best Local Similarity  28.2%;  Pred. No. 3.49e-13;
  Matches    37;  Conservative   32;  Mismatches  59;  Indels   3;  Gaps   3;

Db     60 vvngatvhrsslevheivrasdvrqalrdagdefelryrrafsdltsqlhitpgtayqsf 119
          |   |     :| ::     ::| : |   ||::: ||   |  :  :|: |:  ||: |
Qy     52 VAAPADPEMVTLPLQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENAYEYF 111

Db    120 eqvvnelfhdgvnwgrivaffsfggalcvesvdkemrvlvgrivswmttyltdh-ldpwi 178
            :   ||  | ||||:||:::||  | :   :  :  ::|::   :  ::  | :  ||
Qy    112 TKIATSLFESG-NWGRVVALLGFGYRLALHVYQHGLTGFLGQVTRFVVDFMLHHCIARWI 170

Db    179 qenggwvrtal 189
            : |||| :||
Qy    171 AQRGGWV-AAL 180


RESULT   10
ID   W01020 standard; Protein; 232 AA.
AC   W01020;
DT   18-DEC-1996  (first entry)
DE   Apoptosis-blocking protein Bcl-2 mutant 80-6 (del80-86).
KW   Apoptosis-regulating protein; Bcl-2; oncogene;
KW   adenovirus E1B 19K protein; cell death; cancer; tumour;
KW   immune disorder; diagnosis; therapy; Bip1A; Bip13; Bip5; NIp1;
KW   Nip2; Nip3.
OS   Synthetic.
PN   EP-733706-A2.
PD   25-SEP-1996.
PF   21-MAR-1996; 104542.
PR   21-MAR-1995; US-408095.
PA   (UYSL-) UNIV ST LOUIS.
PI   Chinnadurai G;
DR   WPI; 96-427055/43.
PT   Nucleic acids encoding apoptosis regulating proteins - useful for
PT   diagnosing and treating immune disorders, malignancies, etc.
PS   Example 8; Page 34-35; 60pp; English.
CC   The 80-6 mutant (W01020) of the bcl-2 oncogene product (W01018)
CC   lacks amino acids 80-86 of the native protein.  This and other
CC   Bcl-2 mutants (see also W01019-21) were used in a two hybrid assay
CC   to examine the interactions between Bcl-2 and novel apoptosis-
CC   regulating proteins Nip1, Nip2 and Nip3 (W00997-99).  2 Motifs
CC   (W01003-04) on Bcl-2 were identified that are essential for
CC   interaction with the Nip proteins.  These motifs show homology
```

```
CC   to motifs (W01005-06) identified on the adenovirus E1B 19K
CC   apoptosis-blocking protein (W01010).
SQ   Sequence   232 AA;

  Query Match           16.5%;  Score 256;  DB 19;  Length 232;
  Best Local Similarity  29.7%;  Pred. No. 5.25e-13;
  Matches    47;  Conservative   40;  Mismatches  62;  Indels   9;  Gaps   7;

Db     48 ffssqpghtphp-aasrdpvartsplqtpaapgspvppvvhltlrqagddfsrryrrdfa 106
          |:  |  : :   ||: ||   | ||| ::: |  |    :|::   |||::|||  :|
Qy     40 FYRHQQEQEAEGVAAPADPEMVTLPLQPSSTMG-QVGR--QLAI--IGDDINRRYDSEFQ 94

Db    107 emssqlhltpftargcfatvveelfrdgvnwgrivaffefggvmcvesvnremsplvdni 166
           |  :|: |:  |   |: :   ||  | ||||:||:: ||  : :    : :: ::  :
Qy     95 TMLQHLQPTAENAYEYFTKIATSLFESG-NWGRVVALLGFGYRLALHVYQHGLTGFLGQV 153

Db    167 alwmteylnrh-lhtwiqdnggwdafvel-ygpsmrpl 202
          :  : ::: :| :  || : ||| | ::|  || :  |
Qy    154 TRFVVDFMLHHCIARWIAQRGGWVAALNLGNGPILNVL 191


RESULT   11
ID   R68887 standard; Protein; 233 AA.
AC   R68887;
DT   10-AUG-1995  (first entry)
DE   Human thymus BCL-XL.
KW   BCL-XL; apoptosis; cell death; cancer; neurodegenerative disease;
KW   autoimmune disease; Parkinson disease; amylotrophic lateral sclerosis;
KW   multiple sclerosis.
OS   Homo sapiens.
PN   WO9500642-A.
PD   05-JAN-1995.
PF   22-JUN-1994; U07089.
PR   22-JUN-1993; US-081448.
PA   (ARCH-) ARCH DEV CORP.
PA   (UNMI ) UNIV MICHIGAN.
PI   Boise LH,  Nunez G,  Thompson CB;
DR   WPI; 95-052079/07.
DR   N-PSDB; Q81698.
PT   New poly-nucleotide encoding new poly-peptide(s) that modify
PT   apoptosis - and related vectors, recombinant cells and
PT   antibodies, useful in assay and for control of cell death in e.g.
PT   neuronal cells, lymphocytes and cancers
PS   Claim 3; Page 94; 127pp; English.
CC   This protein may be expressed recombinantly, particularly with pcmV
CC   plasmids as vectors for expression in mammalian cell cultures.  The
CC   protein has particular application in cancer cells (failure of
CC   programmed cell death (PCD)) or neurodegenerative and autoimmune
CC   diseases (premature PCD), e.g. Parkinson's disease, amylotrophic
CC   lateral sclerosis and multiple sclerosis.
SQ   Sequence   233 AA;

  Query Match           16.1%;  Score 250;  DB 13;  Length 233;
  Best Local Similarity  25.5%;  Pred. No. 1.79e-12;
  Matches    36;  Conservative   41;  Mismatches  60;  Indels   4;  Gaps   4;

Db     61 dspavngatghs-ssldarevipmaavkqalreagdefelryrrafsdltsqlhitpgta 119
          :: :| :::     :|  :    |: | : |   ||::: ||   |  :  :|: |:  |
Qy     48 EAEGVAAPADPEMVTLPLQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENA 107

Db    120 yqsfeqvvnelfrdgvnwgrivaffsfggalcvesvdkemqvlvsriaawmatylndh-l 178
          |: |  :   ||  | ||||:||:::||  | :   :  :  :::::  :  ::  | :
Qy    108 YEYFTKIATSLFESG-NWGRVVALLGFGYRLALHVYQHGLTGFLGQVTRFVVDFMLHHCI 166

Db    179 epwiqenggwdtfvelygnna 199
            || : ||| : ::| || :
Qy    167 ARWIAQRGGWVAALNL-GNGP 186


RESULT   12
ID   W05821 standard; Protein; 233 AA.
AC   W05821;
DT   30-MAR-1997  (first entry)
DE   Bcl-XL protein.
KW   Human; bcl-XL; T-lymphocyte; cell death; BH1 domain; BH2 domain;
KW   Bcl-2 homology domain; gene therapy; HIV; AIDS; antisense;
KW   immune disorder; autoimmune disease; graft rejection;
KW   graft-versus-host disease; apoptosis; adoptive immunotherapy.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   Domain          129..148
FT   /note= "BH1 domain"
FT   Domain          180..191
FT   /note= "BH2 domain"
PN   WO9634956-A1.
PD   07-NOV-1996.
PF   02-MAY-1996; U06203.
PR   04-MAY-1995; US-435518.
PR   07-JUN-1995; US-481739.
PA   (ARCH-) ARCH DEV CORP.
PA   (USNA ) US SEC OF NAVY.
PI   June CH,  Thompson CB;
DR   WPI; 96-506159/50.
DR   N-PSDB; T40079.
PT   Inducing or preventing death of T cells by bcl-XL protein regulation
PT   - used to increase survival of HIV infected cells or to
PT   down:regulate immune responses in immune diseases
PS   Disclosure; Page 52-53; 76pp; English.
CC   This is the sequence of a human bcl-XL protein, which protects
CC   T-lymphocytes against cell death.  A splice variant form, bcl-XS,
CC   lacks a stretch of 63 amino acids, and is a dominant negative
CC   regulator of bcl-XL function.  The gene may be modified to
CC   facilitate interaction with costimulatory Bax protein and inhibit
CC   interaction with antagonistic Bad protein, by modification of the
CC   Bcl-2 homology domains BH1 and/or BH2.  The bcl-XL gene may be
CC   introduced into T-cells in vivo or ex vivo via gene transfer using
CC   a vector for HIV infection gene therapy, to augment intracellular
CC   bcl-XL protein levels and protect from cell death.  A corresponding
CC   antisense oligonucleotide or expression vector may be used in gene
CC   therapy of e.g. autoimmune disease, graft rejection or graft-
CC   versus-host disease, to induce cell death (e.g. apoptosis) and
CC   down-regulate the immune response in a T-lymphocyte population.
SQ   Sequence   233 AA;

  Query Match           16.1%;  Score 250;  DB 20;  Length 233;
  Best Local Similarity  25.5%;  Pred. No. 1.79e-12;
  Matches    36;  Conservative   41;  Mismatches  60;  Indels   4;  Gaps   4;

Db     61 dspavngatghs-ssldarevipmaavkqalreagdefelryrrafsdltsqlhitpgta 119
          :: :| :::     :|  :    |: | : |   ||::: ||   |  :  :|: |:  |
Qy     48 EAEGVAAPADPEMVTLPLQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENA 107

Db    120 yqsfeqvvnelfrdgvnwgrivaffsfggalcvesvdkemqvlvsriaawmatylndh-l 178
          |: |  :   ||  | ||||:||:::||  | :   :  :  :::::  :  ::  | :
Qy    108 YEYFTKIATSLFESG-NWGRVVALLGFGYRLALHVYQHGLTGFLGQVTRFVVDFMLHHCI 166

Db    179 epwiqenggwdtfvelygnna 199
            || : ||| : ::| || :
Qy    167 ARWIAQRGGWVAALNL-GNGP 186


RESULT   13
ID   W01019 standard; Protein; 232 AA.
AC   W01019;
DT   18-DEC-1996  (first entry)
DE   Apoptosis-blocking protein Bcl-2 mutant 42-8 (del42-48).
KW   Apoptosis-regulating protein; Bcl-2; oncogene;
KW   adenovirus E1B 19K protein; cell death; cancer; tumour;
KW   immune disorder; diagnosis; therapy; Bip1A; Bip13; Bip5; NIp1;
KW   Nip2; Nip3.
OS   Synthetic.
PN   EP-733706-A2.
PD   25-SEP-1996.
PF   21-MAR-1996; 104542.
PR   21-MAR-1995; US-408095.
PA   (UYSL-) UNIV ST LOUIS.
```

```
PI   Chinnadurai G;
DR   WPI; 96-427055/43.
PT   Nucleic acids encoding apoptosis regulating proteins - useful for
PT   diagnosing and treating immune disorders, malignancies, etc.
PS   Example 8; Page 33-34; 60pp; English.
CC   The 42-8 mutant (W01019) of the bcl-2 oncogene product (W01018)
CC   lacks amino acids 42-48 of the native protein.  This and other
CC   Bcl-2 mutants (see also W01020-21) were used in a two hybrid assay
CC   to examine the interactions between Bcl-2 and novel apoptosis-
CC   regulating proteins Nip1, Nip2 and Nip3 (W00997-99).  The Nip
CC   proteins were unable to interact with mutant 42-8.  The site of
CC   deletion in this mutant corresponds to a motif (see also W01003)
CC   on Bcl-2 essential for interaction with Nip proteins.  A second
CC   binding motif (W01004) of Bcl-2 was also identified, and both
CC   show homology to motifs (W01005-06) found on the 19K protein
CC   (W01010) of adenovirus E1B.
SQ   Sequence   232 AA;

  Query Match           15.8%;  Score 245;  DB 19;  Length 232;
  Best Local Similarity  29.3%;  Pred. No. 4.95e-12;
  Matches    41;  Conservative   36;  Mismatches  57;  Indels   6;  Gaps   5;

Db     67 tpaapgaaagpalspvppvvhlt--lrqagddfsrryrrdfaemssqlhltpftargcfa 124
          :|| |   : | | | : : ::   |   |||::|||  :|  |  :|: |:  |   |:
Qy     54 APADPEMVTLP-LQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENAYEYFT 112

Db    125 tvveelfrdgvnwgrivaffefggvmcvesvnremsplvdnialwmteylnrh-lhtwiq 183
          :   ||  | ||||:||:: ||  : :    : :: ::  ::  : ::: :| :  ||
Qy    113 KIATSLFESG-NWGRVVALLGFGYRLALHVYQHGLTGFLGQVTRFVVDFMLHHCIARWIA 171

Db    184 dnggwdafvel-ygpsmrpl 202
          : ||| | ::|  || :  |
Qy    172 QRGGWVAALNLGNGPILNVL 191


RESULT   14
ID   W01018 standard; Protein; 239 AA.
AC   W01018;
DT   18-DEC-1996  (first entry)
DE   Apoptosis-blocking protein Bcl-2.
KW   Apoptosis-regulating protein; Bcl-2; oncogene;
KW   adenovirus E1B 19K protein; cell death; cancer; tumour;
KW   immune disorder; diagnosis; therapy; Bip1A; Bip13; Bip5; NIp1;
KW   Nip2; Nip3.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   Binding_site    43..51
FT   /label= Binding_motif
FT   /note= "interacts with Bip proteins"
FT   Binding_site    106..112
FT   /label= Binding_motif
FT   /note= "interacts with Bip proteins"
PN   EP-733706-A2.
PD   25-SEP-1996.
PF   21-MAR-1996; 104542.
PR   21-MAR-1995; US-408095.
PA   (UYSL-) UNIV ST LOUIS.
PI   Chinnadurai G;
DR   WPI; 96-427055/43.
PT   Nucleic acids encoding apoptosis regulating proteins - useful for
PT   diagnosing and treating immune disorders, malignancies, etc.
PS   Example 8; Page 32-33; 60pp; English.
CC   The bcl-2 oncogene product (W01018) enhances the survival of
CC   haematopoietic B and T cells by blocking apoptosis induced by
CC   diverse agents.  Its activity is similar to that of the 19K
CC   protein (W01010) of adenovirus E1B.  3 Novel proteins, Bip1A,
CC   Bip13 and Bip5 (W01000-02), that specifically interact with
CC   Bcl-2, have been identified.  Mutational analysis (see also
CC   W01019-21) shows the apoptosis-regulating Nip proteins (see also
CC   W00997-99) associate with Bcl-2 at specific sites (see also
CC   W01003-04) that show homology to motifs (W01005-06) on 19K.
SQ   Sequence   239 AA;

  Query Match           15.8%;  Score 245;  DB 19;  Length 239;
  Best Local Similarity  29.3%;  Pred. No. 4.95e-12;
  Matches    41;  Conservative   36;  Mismatches  57;  Indels   6;  Gaps   5;

Db     74 tpaapgaaagpalspvppvvhlt--lrqagddfsrryrrdfaemssqlhltpftargcfa 131
          :|| |   : | | | : : ::   |   |||::|||  :|  |  :|: |:  |   |:
Qy     54 APADPEMVTLP-LQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENAYEYFT 112

Db    132 tvveelfrdgvnwgrivaffefggvmcvesvnremsplvdnialwmteylnrh-lhtwiq 190
          :   ||  | ||||:||:: ||  : :    : :: ::  ::  : ::: :| :  ||
Qy    113 KIATSLFESG-NWGRVVALLGFGYRLALHVYQHGLTGFLGQVTRFVVDFMLHHCIARWIA 171

Db    191 dnggwdafvel-ygpsmrpl 209
          : ||| | ::|  || :  |
Qy    172 QRGGWVAALNLGNGPILNVL 191


RESULT   15
ID   R71405 standard; protein; 205 AA.
AC   R71405;
DT   30-OCT-1995  (first entry)
DE   Human bcl-2 beta protein.
KW   Human; bcl-2; alpha; beta; proto-oncogene; hematopoietic cell line;
KW   apoptosis; membrane-associated cytoplasmic protein; B cell; T cell;
KW   proliferation; cell cycle progression; Bax; apoptotic cell death;
KW   apoptosis; cytokine;  death repressor; BH1; BH2; cancer therapy;
KW   hyperplasia;  immunodeficiency disease; AIDS; neurodegeneration;
KW   ischaemic cell death.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   Domain          136..155
FT   /label= BH1_domain
FT   /note= "Represents Bax binding site"
FT   Domain          187..192
FT   /label= BH2_domain
FT   /note= "Represents Bax binding site"
PN   WO9505750-A.
PD   02-MAR-1995.
PF   24-AUG-1994; U09701.
PR   26-AUG-1993; US-112208.
PR   25-MAY-1994; US-248819.
PA   (UNIW ) UNIV WASHINGTON.
PI   Korsmeyer SJ;
DR   WPI; 95-106605/14.
PT   Methods for producing and identifying mutant bcl-2 proteins -
PT   that lack death repressor activity and/or lacks binding to Bax.
PS   Disclosure; Page 40; 133pp; English.
CC   The sequences given in R71404-05 represent the human bcl-2 alpha and
CC   beta proteins respectively.  bcl-2 is encoded by a proto-oncogene and
CC   is capable of inhibiting apoptosis in many hematopoietic cell systems.
CC   bcl-2 is a 26 kD membrane-associated cytoplasmic protein and is thought
CC   to function by enhancing the survival of hematopoietic cells of B and T
CC   origins rather than directly promoting proliferation of these cell
CC   types.  bcl-2 has not been shown to directly promote cell cycle
CC   progression nor does it necessarily alter the dose reponse to limiting
CC   concentrations of IL-3.  bcl-2 has been shown to form heterodimers with
CC   a 21 kD protein, Bax.  Overexpressed Bax accelerates apoptotic cell death
CC   induced by cytokine deprivation in an IL-3 dependent cell line, and it
CC   also acts to counter the death repressor activity of bcl-2.  Therefore,
CC   the ratio between bcl-2 and Bax determines cell survival or death
CC   following an apoptotic stimulus.  The invention gives a mutant form of
CC   bcl-2 in which there is at least one amino acid substitution or deletion
CC   in the BH1 or BH2 domains.  This makes the mutant protein substantially
CC   incapable of binding Bax and/or incapable of death repressor activity.
CC   Down regulation of bcl-2 is useful in cancer therapy, controlling
CC   hyperplasias and eliminating self-reactive clones in autoimmunity by
CC   favouring death effector molecules.  Up regulating bcl-2is beneficial in
CC   treatment and diagnosis of immunodeficiency diseases, including AIDS and
CC   neurodegenerative and ischaemic cell death.
SQ   Sequence   205 AA;

  Query Match           15.6%;  Score 243;  DB 13;  Length 205;
  Best Local Similarity  29.7%;  Pred. No. 7.43e-12;
```

Matches 38; Conservative 36; Mismatches 49; Indels 5; Gaps 4;

```
Db      74 tpaapgaaagpalspvppvvhla--lrqagddfsrryrgdfaemssqlhltpftargrfa 131
           :|| |   : | | | : : :::  |   |||::||| ::|  |  :|: |:  |   |:
Qy      54 APADPEMVTLP-LQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENAYEYFT 112

Db     132 tvveelfrdgvnwgrivaffefggvmcvesvnremsplvdnialwmteylnrh-lhtwiq 190
            :   ||  | ||||:||:: ||  : :    : :: ::  ::  : ::: :| :  ||
Qy     113 KIATSLFESG-NWGRVVALLGFGYRLALHVYQHGLTGFLGQVTRFVVDFMLHHCIARWIA 171

Db     191 dnggwvga 198
           : ||||:|
Qy     172 QRGGWVAA 179
```

Search completed: Wed Aug 20 11:12:27 1997
Job time : 25 secs.

****

MPSRCH (TM)

****

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Wed Aug 20 11:12:45 1997; MasPar time 9.69 Seconds
626.144 Million cell updates/sec

Tabular output not generated.

Title: >US-08-320-157-22
Description: (1-210) from US08320157.pep
Perfect Score: 1554
Sequence: 1 MASGQGPGPPRQECGEPALP.....LVVLGVVLLGQFVVRRFFKS 210

Scoring table: PAM 150
Gap 11

Searched: 91006 seqs, 28888923 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: pir51
1:ann1 2:ann2 3:ann3 4:ann4 5:unann1 6:unann2 7:unann3
8:unann4 9:unann5 10:unann6 11:unann7 12:unann8
13:unann9 14:unann10 15:unenc 16:unrev

Statistics: Mean 45.070; Variance 107.726; scale 0.418

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description	Pred. No.
1	1543	99.3	211	13	S58873	Bak protein - human	1.46e-241
2	1511	97.2	211	13	S58875	cdn-2 protein - huma	6.54e-236
3	258	16.6	190	13	A47537	apoptosis regulator	6.91e-22
4	254	16.3	233	14	S51761	BCL-X protein - rat	2.79e-21
5	254	16.3	233	14	I49056	bcl-x long - mouse	2.79e-21
6	250	16.1	233	13	B47537	apoptosis regulator	1.13e-20
7	247	15.9	214	14	I49057	bcl-x transmembrane	3.19e-20
8	246	15.8	216	6	B37332	transforming protein	4.51e-20
9	245	15.8	233	6	A37332	transforming protein	6.38e-20
10	243	15.6	205	2	TVHUB1	transforming protein	1.27e-19
11	239	15.4	206	6	D37332	transforming protein	5.07e-19
12	239	15.4	239	2	TVHUA1	transforming protein	5.07e-19
13	236	15.2	233	14	I67431	BCL-X-Long - rat	1.42e-18
14	234	15.1	199	2	TVMSB1	transforming protein	2.83e-18
15	234	15.1	236	2	TVMSA1	transforming protein	2.83e-18
16	234	15.1	237	6	E37332	transforming protein	2.83e-18
17	231	14.9	236	14	I67432	BCL-2 - rat (fragmen	7.90e-18
18	223	14.4	232	6	S24390	transforming protein	1.21e-16
19	205	13.2	192	14	D47538	programmed cell deat	5.11e-14
20	194	12.5	192	13	A47538	bcl-2-associated pro	1.93e-12
21	189	12.2	218	13	B47538	bcl-2-associated pro	9.88e-12
22	176	11.3	350	13	A47476	BCL2 homolog MCL1 -	6.50e-10
23	174	11.2	133	14	I53295	bax - rat (fragment)	1.23e-09
24	167	10.7	154	14	I58194	gene bcl-2 protein -	1.12e-08
25	165	10.6	143	13	I38921	BAX splice form delt	2.09e-08
26	158	10.2	172	14	I49449	hemopoietic-specific	1.83e-07
27	116	7.5	255	5	S26032	cytochrome-c oxidase	3.59e-02
28	106	6.8	133	4	GPBPP4	gop protein - satell	4.96e-01
29	102	6.6	400	10	S35958	beta-glucosidase - S	1.36e+00
30	102	6.6	636	14	I48689	gene NK10 protein -	1.36e+00
31	103	6.6	890	8	S44150	coat protein - straw	1.06e+00
32	98	6.3	177	13	S54778	NR-13 protein - quai	3.65e+00
33	98	6.3	206	11	B28443	phosphatidyl-N-methy	3.65e+00
34	97	6.2	170	14	I49055	bcl-x short - mouse	4.65e+00
35	96	6.2	470	2	SYPSRA	threonine synthase (	5.92e+00
36	97	6.2	912	14	I58149	metabotropic glutama	4.65e+00
37	96	6.2	2007	3	B43402	myosin heavy chain-B	5.92e+00
38	95	6.1	471	7	A39024	collagen alpha 3(IV)	7.51e+00
39	94	6.0	338	13	I52608	glycoprotein D - hum	9.52e+00
40	94	6.0	338	13	A49640	human blood group an	9.52e+00
41	94	6.0	531	10	S07881	hypothetical protein	9.52e+00
42	94	6.0	535	1	B35182	dimethylaniline mono	9.52e+00
43	94	6.0	822	11	S56801	hypothetical protein	9.52e+00
44	94	6.0	885	11	S22389	acetylglutamate kina	9.52e+00
45	94	6.0	908	14	I49142	metabotropic glutama	9.52e+00

ALIGNMENTS

```
RESULT    1
ENTRY             S58873     #type complete
TITLE             Bak protein - human
ALTERNATE_NAMES   bcl-2 homolog; cdn-1 protein
ORGANISM          #formal_name Homo sapiens #common_name man
DATE              15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change
                    06-Sep-1996
ACCESSIONS        S58873; S58872; S58874
REFERENCE         S58873
   #authors       Chittenden, T.; Harrington, E.A.; O'Connor, R.; Flemington,
                    C.; Lutz, R.J.; Evan, G.I.; Guild, B.C.
   #journal       Nature (1995) 374:733-736
   #title         Induction of apoptosis by the Bcl-2 homologue Bak.
   #accession     S58873
      ##status         preliminary; nucleic acid sequence not shown
      ##molecule_type mRNA
      ##residues       1-211 ##label CHI
      ##cross-references EMBL:U23765
REFERENCE         S58872
   #authors       Farrow, S.N.; White, J.H.M.; Martinou, I.; Raven, T.; Pun,
                    K.T.; Grinham, C.J.; Martinou, J.C.; Brown, R.
   #journal       Nature (1995) 374:731-733
   #title         Cloning of a bcl-2 homologue by interaction with adenovirus
                    E1B 19K.
   #accession     S58872
      ##status         preliminary
      ##molecule_type mRNA
      ##residues       1-211 ##label FAR
      ##cross-references EMBL:X84213
REFERENCE         S58874
   #authors       Kiefer, M.C.; Brauer, M.J.; Powers, V.C.; Wu, J.J.; Umansky,
                    S.R.; Tomei, L.D.; Barr, P.J.
   #journal       Nature (1995) 374:736-739
   #title         Modulation of apoptosis by the widely distributed Bcl-2
                    homologue Bak.
   #accession     S58874
      ##status         preliminary
      ##molecule_type mRNA
      ##residues       1-211 ##label KIE
      ##cross-references EMBL:U16811
GENETICS
   #gene          GDB:BAK-LSB
      ##cross-references GDB:635887
SUMMARY           #length 211  #molecular-weight 23409  #checksum 801
```

```
  Query Match           99.3%;  Score 1543;  DB 13;  Length 211;
  Best Local Similarity  99.5%;  Pred. No. 1.46e-241;
  Matches   210;  Conservative    0;  Mismatches   0;  Indels   1;  Gaps   1;

Db      1 masgqgpgpprqecgepalpsaseeqvaqdteevfrsyvfyrhqqeqeaegvaapadpem 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      1 MASGQGPGPPRQECGEPALPSASEEQVAQDTEEVFRSYVFYRHQQEQEAEGVAAPADPEM 60

Db     61 vtlplqpsstmgqvgrqlaiigddinrrydsefqtmlqhlqptaenayeyftkiatslfe 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     61 VTLPLQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENAYEYFTKIATSLFE 120

Db    121 sginwgrvvallgfgyrlalhvyqhgltgflgqvtrfvvdfmlhhciarwiaqrggwvaa 180
          || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    121 SG-NWGRVVALLGFGYRLALHVYQHGLTGFLGQVTRFVVDFMLHHCIARWIAQRGGWVAA 179

Db    181 lnlgngpilnvlvvlgvvllgqfvvrrffks 211
          |||||||||||||||||||||||||||||||
Qy    180 LNLGNGPILNVLVVLGVVLLGQFVVRRFFKS 210


RESULT    2
ENTRY             S58875     #type complete
TITLE             cdn-2 protein - human
ORGANISM          #formal_name Homo sapiens #common_name man
DATE              15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change
                    01-Mar-1996
ACCESSIONS        S58875
REFERENCE         S58874
   #authors       Kiefer, M.C.; Brauer, M.J.; Powers, V.C.; Wu, J.J.; Umansky,
                    S.R.; Tomei, L.D.; Barr, P.J.
   #journal       Nature (1995) 374:736-739
   #title         Modulation of apoptosis by the widely distributed Bcl-2
                    homologue Bak.
   #accession     S58875
      ##status         preliminary; nucleic acid sequence not shown;
                         translation not shown
      ##molecule_type DNA
      ##residues       1-211 ##label KIE
      ##cross-references EMBL:U16812
      ##note           the nucleotide sequence was submitted to the EMBL Data
                         Library, November 1994
SUMMARY           #length 211  #molecular-weight 23411  #checksum 9485

  Query Match           97.2%;  Score 1511;  DB 13;  Length 211;
  Best Local Similarity  96.7%;  Pred. No. 6.54e-236;
  Matches   204;  Conservative    5;  Mismatches   1;  Indels   1;  Gaps   1;

Db      1 masgqgpgpprqecgepalpsaseeqvaqdteevfrsyvfyhhqqeqeaegaaapadpem 60
          ||||||||||||||||||||||||||||||||||||||:|||||||| ||||||||
Qy      1 MASGQGPGPPRQECGEPALPSASEEQVAQDTEEVFRSYVFYRHQQEQEAEGVAAPADPEM 60

Db     61 vtlplqpsstmgqvgrqlaiigddinrrydsefqtmlqhlqptaenayeyftkiasslfe 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||
Qy     61 VTLPLQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENAYEYFTKIATSLFE 120

Db    121 sginwgrvvallgfsyrlalhiyqrgltgflgqvtrfvvdfmlhhciarwiaqrggwvaa 180
          || |||||||||||:||||||:||:|||||||||||||||||||||||||||||||||||
Qy    121 SG-NWGRVVALLGFGYRLALHVYQHGLTGFLGQVTRFVVDFMLHHCIARWIAQRGGWVAA 179

Db    181 lnlgngpilnvlvvlgvvllgqfvvrrffks 211
          |||||||||||||||||||||||||||||||
Qy    180 LNLGNGPILNVLVVLGVVLLGQFVVRRFFKS 210


RESULT    3
ENTRY             A47537     #type complete
TITLE             apoptosis regulator bcl-x - chicken
ORGANISM          #formal_name Gallus gallus #common_name chicken
DATE              03-May-1994 #sequence_revision 03-May-1994 #text_change
                    03-May-1994
ACCESSIONS        A47537
REFERENCE         A47537
   #authors       Boise, L.H.; Gonzalez-Garcia, M.; Postema, C.E.; Ding, L.;
                    Lindsten, T.; Turka, L.A.; Mao, X.; Nunez, G.; Thompson,
                    C.B.
   #journal       Cell (1993) 74:597-608
   #title         bcl-x, a bcl-2-related gene that functions as a dominant
                    regulator of apoptotic cell death.
   #accession     A47537
      ##status         preliminary
      ##molecule_type DNA
      ##residues       1-190 ##label BOI
      ##cross-references GB:L20120
SUMMARY           #length 190  #molecular-weight 21467  #checksum 5509

  Query Match           16.6%;  Score 258;  DB 13;  Length 190;
  Best Local Similarity  28.2%;  Pred. No. 6.91e-22;
  Matches    37;  Conservative   32;  Mismatches  59;  Indels   3;  Gaps   3;

Db     60 vvngatvhrsslevheivrasdvrqalrdagdefelryrrafsdltsqlhitpgtayqsf 119
          |  |     :| ::     ::| : |   ||::: ||   |  :  :|: |:  ||: |
Qy     52 VAAPADPEMVTLPLQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENAYEYF 111

Db    120 eqvvnelfhdgvnwgrivaffsfggalcvesvdkemrvlvgrivswmttyltdh-ldpwi 178
          :   ||  | ||||:||:::||  | :   :  :  ::|::   :  ::  | :  ||
Qy    112 TKIATSLFESG-NWGRVVALLGFGYRLALHVYQHGLTGFLGQVTRFVVDFMLHHCIARWI 170

Db    179 qenggwvrtal 189
          : |||| :||
Qy    171 AQRGGWV-AAL 180


RESULT    4
ENTRY             S51761     #type complete
TITLE             BCL-X protein - rat
ORGANISM          #formal_name Rattus norvegicus #common_name Norway rat
DATE              07-May-1995 #sequence_revision 01-Sep-1995 #text_change
                    01-Dec-1995
ACCESSIONS        S51761; S51762
REFERENCE         S51761
   #authors       Michaelidis, T.M.
   #submission    submitted to the EMBL Data Library, November 1994
   #accession     S51761
      ##status         preliminary
      ##molecule_type DNA
      ##residues       1-233 ##label MIC
      ##cross-references EMBL:X82537
REFERENCE         S51761
   #authors       Michaelidis, T.M.
   #submission    submitted to the EMBL Data Library, November 1994
   #accession     S51762
      ##status         preliminary
      ##molecule_type DNA
      ##residues       1-125,189-233 ##label MI2
      ##cross-references EMBL:X82537
GENETICS
   #introns       125/3
SUMMARY           #length 233  #molecular-weight 26130  #checksum 6378

  Query Match           16.3%;  Score 254;  DB 14;  Length 233;
  Best Local Similarity  25.5%;  Pred. No. 2.79e-21;
  Matches    36;  Conservative   41;  Mismatches  60;  Indels   4;  Gaps   4;

Db     61 dspavngatghs-ssldarevipmaavkqalreagdefelryrrafsdltsqlhitpgta 119
          :: :| :::     :|  :    |: | : |   ||::: ||   |  :  :|: |:  |
Qy     48 EAEGVAAPADPEMVTLPLQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENA 107

Db    120 yqsfeqvvnelfrdgvnwgrivaffsfggalcvesvdkemqvlvsriaswmatylndh-l 178
          |: |  :   ||  | ||||:||:::||  | :   :  :  ::::::  :  ::  | :
Qy    108 YEYFTKIATSLFESG-NWGRVVALLGFGYRLALHVYQHGLTGFLGQVTRFVVDFMLHHCI 166

Db    179 epwiqenggwdtfvdlygnna 199
          || : ||| : ::| || :
Qy    167 ARWIAQRGGWVAALNL-GNGP 186
```

```
RESULT    5
ENTRY             I49056     #type complete
TITLE             bcl-x long - mouse
ORGANISM          #formal_name Mus musculus #common_name house mouse
DATE              02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change
                    15-Oct-1996
ACCESSIONS        I49056; S52866
REFERENCE         I49055
   #authors       Fang, W.; Rivard, J.J.; Mueller, D.L.; Behrens, T.W.
   #journal       J. Immunol. (1994) 153:4388-4398
   #title         Cloning and molecular characterization of mouse bcl-x in B
                    and T lymphocytes.
   #cross-references MUID:95052604
   #accession     I49056
      ##status         preliminary; translated from GB/EMBL/DDBJ
      ##molecule_type mRNA
      ##residues       1-233 ##label RES
      ##cross-references EMBL:U10101; NID:g506647; CDS_PID:g506648
REFERENCE         S52866
   #authors       Kamesaki, H.; Michaud, G.Y.; Takatsu, K.; Okuma, M.
   #submission    submitted to the EMBL Data Library, November 1994
   #description   IL-5 inhibits anti-IgM-induced apoptosis in an immature B
                    cell line through inductin of bcl-Xl.
   #accession     S52866
      ##status         preliminary
      ##molecule_type mRNA
      ##residues       1-233 ##label KAM
      ##cross-references EMBL:X83574
SUMMARY           #length 233  #molecular-weight 26132  #checksum 5739

  Query Match           16.3%;  Score 254;  DB 14;  Length 233;
  Best Local Similarity  25.5%;  Pred. No. 2.79e-21;
  Matches    36;  Conservative   41;  Mismatches  60;  Indels   4;  Gaps   4;

Db      61 dspavngatghs-ssldarevipmaavkqalreagdefelryrrafsdltsqlhitpgta 119
           :: :| :::     :|  :    |: | : |   ||::: ||   |  :  :|: |:  |
Qy      48 EAEGVAAPADPEMVTLPLQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENA 107

Db     120 yqsfeqvvnelfrdgvnwgrivaffsfggalcvesvdkemqvlvsriaswmatylndh-l 178
           |: |  :   ||  | ||||:||:::|| | :   :  :  :::::: :  ::  | :
Qy     108 YEYFTKIATSLFESG-NWGRVVALLGFGYRLALHVYQHGLTGFLGQVTRFVVDFMLHHCI 166

Db     179 epwiqenggwdtfvdlygnna 199
             || : ||| : ::| || :
Qy     167 ARWIAQRGGWVAALNL-GNGP 186


RESULT    6
ENTRY             B47537     #type complete
TITLE             apoptosis regulator bcl-xL - human
ALTERNATE_NAMES   bcl-2-related protein
CONTAINS          apoptosis regulator bcl-xS
ORGANISM          #formal_name Homo sapiens #common_name man
DATE              16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change
                    06-Sep-1996
ACCESSIONS        B47537; C47537
REFERENCE         A47537
   #authors       Boise, L.H.; Gonzalez-Garcia, M.; Postema, C.E.; Ding, L.;
                    Lindsten, T.; Turka, L.A.; Mao, X.; Nunez, G.; Thompson,
                    C.B.
   #journal       Cell (1993) 74:597-608
   #title         bcl-x, a bcl-2-related gene that functions as a dominant
                    regulator of apoptotic cell death.
   #accession     B47537
      ##status         nucleic acid sequence not shown; translated from
                         GB/EMBL/DDBJ
      ##molecule_type mRNA
      ##residues       1-233 ##label BOI
      ##cross-references GB:L20121; CDS_PID:Q07817
   #accession     C47537
      ##status         nucleic acid sequence not shown; translated from
                         GB/EMBL/DDBJ
      ##molecule_type mRNA
      ##residues       1-69,'G',71-125,189-233 ##label BO2
      ##cross-references GB:L20122; CDS_PID:g623237
GENETICS
   #gene          GDB:BCL2L
      ##cross-references GDB:228079
KEYWORDS          alternative splicing; apoptosis
FEATURE
   1-233               #product apoptosis regulator bcl-xL #status predicted
                         #label MAT\
   1-125,189-233       #product apoptosis regulator bcl-xS #status predicted
                         #label MA2
SUMMARY           #length 233  #molecular-weight 26063  #checksum 5340

  Query Match           16.1%;  Score 250;  DB 13;  Length 233;
  Best Local Similarity  25.5%;  Pred. No. 1.13e-20;
  Matches    36;  Conservative   41;  Mismatches  60;  Indels   4;  Gaps   4;

Db      61 dspavngatahs-ssldarevipmaavkqalreagdefelryrrafsdltsqlhitpgta 119
           :: :| :::     :|  :    |: | : |   ||::: ||   |  :  :|: |:  |
Qy      48 EAEGVAAPADPEMVTLPLQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENA 107

Db     120 yqsfeqvvnelfrdgvnwgrivaffsfggalcvesvdkemqvlvsriaawmatylndh-l 178
           |: |  :   ||  | ||||:||:::|| | :   :  :  :::::: :  ::  | :
Qy     108 YEYFTKIATSLFESG-NWGRVVALLGFGYRLALHVYQHGLTGFLGQVTRFVVDFMLHHCI 166

Db     179 epwiqenggwdtfvelygnna 199
             || : ||| : ::| || :
Qy     167 ARWIAQRGGWVAALNL-GNGP 186


RESULT    7
ENTRY             I49057     #type complete
TITLE             bcl-x transmembrane deleted - mouse
ORGANISM          #formal_name Mus musculus #common_name house mouse
DATE              02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change
                    02-Jul-1996
ACCESSIONS        I49057
REFERENCE         I49055
   #authors       Fang, W.; Rivard, J.J.; Mueller, D.L.; Behrens, T.W.
   #journal       J. Immunol. (1994) 153:4388-4398
   #title         Cloning and molecular characterization of mouse bcl-x in B
                    and T lymphocytes.
   #cross-references MUID:95052604
   #accession     I49057
      ##status         preliminary; translated from GB/EMBL/DDBJ
      ##molecule_type mRNA
      ##residues       1-214 ##label RES
      ##cross-references EMBL:U10102; NID:g506649; CDS_PID:g506650
GENETICS
   #note          gene name bcl-x long
SUMMARY           #length 214  #molecular-weight 23900  #checksum 9730

  Query Match           15.9%;  Score 247;  DB 14;  Length 214;
  Best Local Similarity  25.4%;  Pred. No. 3.19e-20;
  Matches    33;  Conservative   37;  Mismatches  57;  Indels   3;  Gaps   3;

Db      61 dspavngatghs-ssldarevipmaavkqalreagdefelryrrafsdltsqlhitpgta 119
           :: :| :::     :|  :    |: | : |   ||::: ||   |  :  :|: |:  |
Qy      48 EAEGVAAPADPEMVTLPLQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENA 107

Db     120 yqsfeqvvnelfrdgvnwgrivaffsfggalcvesvdkemqvlvsriaswmatylndh-l 178
           |: |  :   ||  | ||||:||:::|| | :   :  :  :::::: :  ::  | :
Qy     108 YEYFTKIATSLFESG-NWGRVVALLGFGYRLALHVYQHGLTGFLGQVTRFVVDFMLHHCI 166

Db     179 epwiqenggw 188
             || : |||
Qy     167 ARWIAQRGGW 176


RESULT    8
ENTRY             B37332     #type complete
TITLE             transforming protein (bcl-2-beta) - chicken
```

```
ORGANISM          #formal_name Gallus gallus #common_name chicken
DATE              03-Mar-1993 #sequence_revision 03-Mar-1993 #text_change
                    15-Jun-1996
ACCESSIONS        B37332
REFERENCE         A37332
   #authors       Eguchi, Y.; Ewert, D.L.; Tsujimoto, Y.
   #journal       Nucleic Acids Res. (1992) 20:4187-4192
   #title         Isolation and characterization of the chicken bcl-2 gene:
                    expression in a variety of tissues including lymphoid and
                    neuronal organs in adult and embryo.
   #accession     B37332
      ##status         preliminary; nucleic acid sequence not shown; not
                         compared with conceptual translation
      ##molecule_type DNA
      ##residues      1-216 ##label EGU
      ##cross-references EMBL:D11382
CLASSIFICATION    #superfamily bcl transforming protein
SUMMARY           #length 216  #molecular-weight 23492  #checksum 3596

  Query Match            15.8%;  Score 246;  DB 6;  Length 216;
  Best Local Similarity  28.5%;  Pred. No. 4.51e-20;
  Matches    39;  Conservative   37;  Mismatches  56;  Indels   5;  Gaps   5;

Db      60 hhrpeppgsaaasevppae-glrpapp-g-vhlalrqagdefsrryqrdfaqmsgqlhlt 116
           ::  |: | ||:::      |:|::  | |   |   ||:::|||: :|  |  :|: |
Qy      44 QQEQEAEGVAAPADPEMVTLPLQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPT 103

Db     117 pftahgrfvavveelfrdgvnwgrivaffefggvmcvesvnremsplvdniatwmteyln 176
           :  |   |  :   ||  | ||||:||:: ||  : :    : :: ::  ::  : :::
Qy     104 AENAYEYFTKIATSLFESG-NWGRVVALLGFGYRLALHVYQHGLTGFLGQVTRFVVDFML 162

Db     177 rh-lhnwiqdnggwvra 192
           :| :  || : |||| |
Qy     163 HHCIARWIAQRGGWVAA 179


RESULT    9
ENTRY             A37332     #type complete
TITLE             transforming protein (bcl-2-alpha) - chicken
ORGANISM          #formal_name Gallus gallus #common_name chicken
DATE              03-Mar-1993 #sequence_revision 03-Mar-1993 #text_change
                    12-Apr-1995
ACCESSIONS        A37332; S35453
REFERENCE         A37332
   #authors       Eguchi, Y.; Ewert, D.L.; Tsujimoto, Y.
   #journal       Nucleic Acids Res. (1992) 20:4187-4192
   #title         Isolation and characterization of the chicken bcl-2 gene:
                    expression in a variety of tissues including lymphoid and
                    neuronal organs in adult and embryo.
   #accession     A37332
      ##status         nucleic acid sequence not shown
      ##molecule_type DNA
      ##residues      1-233 ##label EGU
      ##cross-references EMBL:D11381
GENETICS
   #introns       189/3
CLASSIFICATION    #superfamily bcl transforming protein
KEYWORDS          transforming protein
SUMMARY           #length 233  #molecular-weight 25687  #checksum 99

  Query Match            15.8%;  Score 245;  DB 6;  Length 233;
  Best Local Similarity  28.5%;  Pred. No. 6.38e-20;
  Matches    41;  Conservative   40;  Mismatches  57;  Indels   6;  Gaps   6;

Db      60 hhrpeppgsaaasevppae-glrpapp-g-vhlalrqagdefsrryqrdfaqmsgqlhlt 116
           ::  |: | ||:::      |:|::  | |   |   ||:::|||: :|  |  :|: |
Qy      44 QQEQEAEGVAAPADPEMVTLPLQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPT 103

Db     117 pftahgrfvavveelfrdgvnwgrivaffefggvmcvesvnremsplvdniatwmteyln 176
           :  |   |  :   ||  | ||||:||:: ||  : :    : :: ::  ::  : :::
Qy     104 AENAYEYFTKIATSLFESG-NWGRVVALLGFGYRLALHVYQHGLTGFLGQVTRFVVDFML 162

Db     177 rh-lhnwiqdnggwdafvelygns 199
           :| :  || : |||| | ::| ||:
Qy     163 HHCIARWIAQRGGWVAALNL-GNG 185


RESULT   10
ENTRY             TVHUB1     #type complete
TITLE             transforming protein bcl-2-beta - human
ORGANISM          #formal_name Homo sapiens #common_name man
DATE              31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change
                    06-Sep-1996
ACCESSIONS        B29409
REFERENCE         A29409
   #authors       Tsujimoto, Y.; Croce, C.M.
   #journal       Proc. Natl. Acad. Sci. U.S.A. (1986) 83:5214-5218
   #title         Analysis of the structure, transcripts, and protein products
                    of bcl-2, the gene involved in human follicular lymphoma.
   #cross-references MUID:86259760
   #accession     B29409
      ##molecule_type mRNA
      ##residues      1-205 ##label TSU
GENETICS
   #gene          GDB:BCL2
      ##cross-references GDB:119031
   #map_position 18q21.33-18q21.33
CLASSIFICATION    #superfamily bcl transforming protein
KEYWORDS          alternative splicing; B-cell lymphoma; follicular lymphoma;
                    transforming protein
SUMMARY           #length 205  #molecular-weight 22182  #checksum 1183

  Query Match            15.6%;  Score 243;  DB 2;  Length 205;
  Best Local Similarity  29.7%;  Pred. No. 1.27e-19;
  Matches    38;  Conservative   36;  Mismatches  49;  Indels   5;  Gaps   4;

Db      74 tpaapgaaagpalspvppvvhla--lrqagddfsrryrgdfaemssqlhltpftargrfa 131
           :|| |   : | | | : : :::  |   |||::||| ::|  |  :|: |:  |   |:
Qy      54 APADPEMVTLP-LQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENAYEYFT 112

Db     132 tvveelfrdgvnwgrivaffefggvmcvesvnremsplvdnialwmteylnrh-lhtwiq 190
           :   ||  | ||||:||:: ||  : :    : :: ::  ::  : ::: :| :  ||
Qy     113 KIATSLFESG-NWGRVVALLGFGYRLALHVYQHGLTGFLGQVTRFVVDFMLHHCIARWIA 171

Db     191 dnggwvga 198
           : ||||:|
Qy     172 QRGGWVAA 179


RESULT   11
ENTRY             D37332     #type complete
TITLE             transforming protein (bcl-2-beta) - human
ORGANISM          #formal_name Homo sapiens #common_name man
DATE              03-Mar-1993 #sequence_revision 03-Mar-1993 #text_change
                    15-Jun-1996
ACCESSIONS        D37332
REFERENCE         A37332
   #authors       Eguchi, Y.; Ewert, D.L.; Tsujimoto, Y.
   #journal       Nucleic Acids Res. (1992) 20:4187-4192
   #title         Isolation and characterization of the chicken bcl-2 gene:
                    expression in a variety of tissues including lymphoid and
                    neuronal organs in adult and embryo.
   #accession     D37332
      ##status         preliminary; nucleic acid sequence not shown; not
                         compared with conceptual translation
      ##molecule_type DNA
      ##residues      1-206 ##label EGU
CLASSIFICATION    #superfamily bcl transforming protein
SUMMARY           #length 206  #molecular-weight 22440  #checksum 5581

  Query Match            15.4%;  Score 239;  DB 6;  Length 206;
  Best Local Similarity  29.7%;  Pred. No. 5.07e-19;
  Matches    38;  Conservative   34;  Mismatches  51;  Indels   5;  Gaps   4;

Db      75 tpaapgaaagpalspvppvvhlt--lrqagddfsrryrrdfaemssqlhltpftargrfa 132
           :|| |   : | | | : : ::   |   |||::|||  :|  |  :|: |:  |   |:
```

```
Qy      54 APADPEMVTLP-LQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENAYEYFT 112

Db     133 tvveelfrdgvnwgrivaffefggvmcvesvnremsplvdnialwmteylnrh-lhtwiq 191
Qy     113 KIATSLFESG-NWGRVVALLGFGYRLALHVYQHGLTGFLGQVTRFVVDFMLHHCIARWIA 171

Db     192 dnggwvga 199
Qy     172 QRGGWVAA 179


RESULT   12
ENTRY            TVHUA1     #type complete
TITLE            transforming protein bcl-2, splice form alpha - human
ORGANISM         #formal_name Homo sapiens #common_name man
DATE             31-Dec-1988 #sequence_revision 07-Jun-1996 #text_change
                   06-Sep-1996
ACCESSIONS       C37332; A29409; S02452; A24428; A27622; B27622
REFERENCE        A37332
   #authors      Eguchi, Y.; Ewert, D.L.; Tsujimoto, Y.
   #journal      Nucleic Acids Res. (1992) 20:4187-4192
   #title        Isolation and characterization of the chicken bcl-2 gene:
                   expression in a variety of tissues including lymphoid and
                   neuronal organs in adult and embryo.
   #accession    C37332
      ##status        nucleic acid sequence not shown; not compared with
                        conceptual translation
      ##molecule_type DNA
      ##residues      1-239 ##label EGU
      ##note          this report is a correction
REFERENCE        A29409
   #authors      Tsujimoto, Y.; Croce, C.M.
   #journal      Proc. Natl. Acad. Sci. U.S.A. (1986) 83:5214-5218
   #title        Analysis of the structure, transcripts, and protein products
                   of bcl-2, the gene involved in human follicular lymphoma.
   #cross-references MUID:86259760
   #accession    A29409
      ##molecule_type mRNA
      ##residues      1-95,'A',97-109,'G',111-236,'S',238-239 ##label TSU
      ##note          this sequence has been corrected in reference A37332
REFERENCE        S02452
   #authors      Seto, M.; Jaeger, U.; Hockett, R.D.; Graninger, W.; Bennett,
                   S.; Goldman, P.; Korsmeyer, S.J.
   #journal      EMBO J. (1988) 7:123-131
   #title        Alternative promoters and exons, somatic mutation and
                   deregulation of the Bcl-2--Ig fusion gene in lymphoma.
   #cross-references MUID:88196071
   #accession    S02452
      ##molecule_type mRNA
      ##residues      1-239 ##label SET
REFERENCE        A24428
   #authors      Cleary, M.L.; Smith, S.D.; Sklar, J.
   #journal      Cell (1986) 47:19-28
   #title        Cloning and structural analysis of cDNAs for bcl-2 and a
                   hybrid bcl-2/immunoglobulin transcript resulting from the t
                   (14;18) translocation.
   #cross-references MUID:87002488
   #accession    A24428
      ##molecule_type mRNA
      ##residues      1-58,'T',60-116,'R',118-239 ##label CLE
REFERENCE        A27622
   #authors      Hua, C.; Zorn, S.; Jensen, J.P.; Coupland, R.W.; Ko, H.S.;
                   Wright, J.J.; Bakhshi, A.
   #journal      Oncogene Res. (1988) 2:263-275
   #title        Consequences of the t(14;18) chromosomal translocation in
                   follicular lymphoma: deregulated expression of a chimeric
                   and mutated BCL-2 gene.
   #cross-references MUID:88217344
   #accession    A27622
      ##molecule_type mRNA
      ##residues      1-58,'T',60-239 ##label HUA
   #accession    B27622
      ##molecule_type DNA
      ##residues      1-6,'S',8-58,'T',60-128,'C',130-239 ##label HUA2
      ##note          the sequence was determined from the germline gene
COMMENT          Constitutive expression of BCL2 following t(14:18) chromosomal
                   translocation is typically found in follicular lymphoma.
GENETICS
   #gene         GDB:BCL2
      ##cross-references GDB:119031
   #map_position 18q21.33-18q21.33
FUNCTION
   #description  blocks apoptosis in hematopoietic cells
CLASSIFICATION   #superfamily bcl transforming protein
KEYWORDS         alternative splicing; B-cell lymphoma; follicular lymphoma;
                   proto-oncogene; transforming protein
SUMMARY          #length 239  #molecular-weight 26266  #checksum 8323

  Query Match           15.4%;  Score 239;  DB 2;  Length 239;
  Best Local Similarity  29.3%;  Pred. No. 5.07e-19;
  Matches    41;  Conservative   36;  Mismatches  57;  Indels   6;  Gaps   5;

Db      74 tpaapgaaagpalspvppvvhlt--lrqagddfsrryrrdfaemssqlhltpftargrfa 131
Qy      54 APADPEMVTLP-LQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENAYEYFT 112

Db     132 tvveelfrdgvnwgrivaffefggvmcvesvnremsplvdnialwmteylnrh-lhtwiq 190
Qy     113 KIATSLFESG-NWGRVVALLGFGYRLALHVYQHGLTGFLGQVTRFVVDFMLHHCIARWIA 171

Db     191 dnggwdafvel-ygpsmrpl 209
Qy     172 QRGGWVAALNLGNGPILNVL 191


RESULT   13
ENTRY            I67431     #type complete
TITLE            BCL-X-Long - rat
ORGANISM         #formal_name Rattus norvegicus #common_name Norway rat
DATE             26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change
                   26-Jul-1996
ACCESSIONS       I67431
REFERENCE        I53295
   #authors      Tilly, J.L.; Tilly, K.I.; Kenton, M.L.; Johnson, A.L.
   #journal      Endocrinology (1995) 136:232-241
   #title        Expression of members of the bcl-2 gene family in the
                   immature rat ovary: equine chorionic gonadotropin-mediated
                   inhibition of granulosa cell apoptosis is associated with
                   decreased bax and constitutive bcl-2 and bcl-xlong
                   messenger ribonucleic acid levels.
   #cross-references MUID:95129487
   #accession    I67431
      ##status        preliminary; translated from GB/EMBL/DDBJ
      ##molecule_type mRNA
      ##residues      1-233 ##label RES
      ##cross-references EMBL:U34963; NID:g1004376; CDS_PID:g1004377
SUMMARY          #length 233  #molecular-weight 26122  #checksum 8310

  Query Match           15.2%;  Score 236;  DB 14;  Length 233;
  Best Local Similarity  28.4%;  Pred. No. 1.42e-18;
  Matches    33;  Conservative   30;  Mismatches  50;  Indels   3;  Gaps   3;

Db      83 maavkqalreagdefelryrrafsdltsqlhitpgtvyqsfeqvvnelfrdgvnwgriva 142
Qy      71 MGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENAYEYFTKIATSLFESG-NWGRVVA 129

Db     143 sssfggalcvesvdkemqvlvsriaswmatylndh-lepwiqenggwdtfvdlygn 197
Qy     130 LLGFGYRLALHVYQHGLTGFLGQVTRFVVDFMLHHCIARWIAQRGGWVAALNL-GN 184


RESULT   14
ENTRY            TVMSB1     #type complete
TITLE            transforming protein bcl-2-beta - mouse
ORGANISM         #formal_name Mus musculus #common_name house mouse
DATE             31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change
```

```
                  02-Jun-1994
ACCESSIONS        B25960
REFERENCE         A90893
   #authors       Negrini, M.; Silini, E.; Kozak, C.; Tsujimoto, Y.; Croce,
                    C.M.
   #journal       Cell (1987) 49:455-463
   #title         Molecular analysis of mbcl-2: structure and expression of the
                    murine gene homologous to the human gene involved in
                    follicular lymphoma.
   #cross-references MUID:87187643
   #accession     B25960
      ##molecule_type DNA
      ##residues      1-199 ##label NEG
GENETICS
   #gene          BCL2
CLASSIFICATION    #superfamily bcl transforming protein
KEYWORDS          alternative splicing; transforming protein
SUMMARY           #length 199  #molecular-weight 22299  #checksum 7397

  Query Match           15.1%;  Score 234;  DB 2;  Length 199;
  Best Local Similarity  31.1%;  Pred. No. 2.83e-18;
  Matches    32;  Conservative   28;  Mismatches  41;  Indels   2;  Gaps   2;

Db     94 lrragddfsrryrrdfaemssqlhltpftargrfatvveelfrdgvnwgrivaffefggv 153
          |   |||::|||  :|  |  :|: |:  |   |: :   ||  | ||||:||:: ||
Qy     78 LAIIGDDINRRYDSEFQTMLQHLQPTAENAYEYFTKIATSLFESG-NWGRVVALLGFGYR 136

Db    154 mcvesvnremsplvdnialwmteylnrh-lhtwiqdnggwvga 195
          : :    : :: ::  ::  : ::: :| :  || : ||||:|
Qy    137 LALHVYQHGLTGFLGQVTRFVVDFMLHHCIARWIAQRGGWVAA 179


RESULT   15
ENTRY             TVMSA1     #type complete
TITLE             transforming protein bcl-2-alpha - mouse
ORGANISM          #formal_name Mus musculus #common_name house mouse
DATE              31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change
                    02-Jun-1994
ACCESSIONS        A25960
REFERENCE         A90893
   #authors       Negrini, M.; Silini, E.; Kozak, C.; Tsujimoto, Y.; Croce,
                    C.M.
   #journal       Cell (1987) 49:455-463
   #title         Molecular analysis of mbcl-2: structure and expression of the
                    murine gene homologous to the human gene involved in
                    follicular lymphoma.
   #cross-references MUID:87187643
   #accession     A25960
      ##molecule_type DNA
      ##residues      1-236 ##label NEG
GENETICS
   #gene          BCL2
   #introns       192/3
CLASSIFICATION    #superfamily bcl transforming protein
KEYWORDS          alternative splicing; transforming protein
SUMMARY           #length 236  #molecular-weight 26524  #checksum 6709

  Query Match           15.1%;  Score 234;  DB 2;  Length 236;
  Best Local Similarity  30.4%;  Pred. No. 2.83e-18;
  Matches    35;  Conservative   30;  Mismatches  47;  Indels   3;  Gaps   3;

Db     94 lrragddfsrryrrdfaemssqlhltpftargrfatvveelfrdgvnwgrivaffefggv 153
          |   |||::|||  :|  |  :|: |:  |   |: :   ||  | ||||:||:: ||
Qy     78 LAIIGDDINRRYDSEFQTMLQHLQPTAENAYEYFTKIATSLFESG-NWGRVVALLGFGYR 136

Db    154 mcvesvnremsplvdnialwmteylnrh-lhtwiqdnggwdafvel-ygpsmrpl 206
          : :    : :: ::  ::  : ::: :| :  || : ||| | ::|  || :  |
Qy    137 LALHVYQHGLTGFLGQVTRFVVDFMLHHCIARWIAQRGGWVAALNLGNGPILNVL 191


Search completed: Wed Aug 20 11:13:22 1997
Job time : 37 secs.
```

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Wed Aug 20 11:13:40 1997; MasPar time 6.35 Seconds
701.611 Million cell updates/sec

Tabular output not generated.

Title: >US-08-320-157-22
Description: (1-210) from US08320157.pep
Perfect Score: 1554
Sequence: 1 MASGQGPGPPRQECGEPALP.....LVVLGVVLLGQFVVRRFFKS 210

Scoring table: PAM 150
Gap 11

Searched: 59021 seqs, 21210388 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: swiss-prot34
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11

Statistics: Mean 46.936; Variance 89.738; scale 0.523

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description	Pred. No.
1	258	16.6	190	1	BCLX_CHICK	APOPTOSIS REGULATOR B	1.75e-27
2	254	16.3	233	1	BCLX_RAT	APOPTOSIS REGULATOR B	9.85e-27
3	250	16.1	233	1	BCLX_HUMAN	APOPTOSIS REGULATOR B	5.51e-26
4	245	15.8	233	1	BCL2_CHICK	APOPTOSIS REGULATOR B	4.70e-25
5	239	15.4	205	1	BC2B_HUMAN	PROTEIN BCL-2-BETA.	6.08e-24
6	239	15.4	239	1	BC2A_HUMAN	PROTEIN BCL-2-ALPHA.	6.08e-24
7	234	15.1	199	1	BC2B_MOUSE	PROTEIN BCL-2-BETA.	5.08e-23
8	234	15.1	236	1	BC2A_MOUSE	PROTEIN BCL-2 ALPHA.	5.08e-23
9	232	14.9	236	1	BCL2_RAT	PROTEIN BCL-2 ALPHA.	1.18e-22
10	206	13.3	192	1	BAXA_MOUSE	APOPTOSIS REGULATOR B	5.98e-18
11	194	12.5	192	1	BAXA_HUMAN	APOPTOSIS REGULATOR B	7.87e-16
12	189	12.2	218	1	BAXB_HUMAN	APOPTOSIS REGULATOR B	5.86e-15
13	176	11.3	350	6	MCL1_HUMAN	INDUCED MYELOID LEUKE	1.00e-12
14	165	10.6	143	1	BAXD_HUMAN	BAX PROTEIN, CYTOPLAS	7.08e-11
15	158	10.2	172	5	HSA1_MOUSE	HEMOPOIETIC-SPECIFIC	1.01e-09
16	126	8.1	179	3	EAR_ASFE4	APOPTOSIS REGULATOR B	1.00e-04
17	121	7.8	179	3	EAR_ASFB7	APOPTOSIS REGULATOR B	5.41e-04
18	118	7.6	179	3	EAR_ASFM2	APOPTOSIS REGULATOR B	1.46e-03
19	116	7.5	255	2	COX3_CAEEL	CYTOCHROME C OXIDASE	2.80e-03
20	107	6.9	626	5	HTPG_BACSU	HEAT SHOCK PROTEIN HT	4.87e-02
21	106	6.8	133	4	GOP_BPP4	GOP PROTEIN.	6.63e-02
22	98	6.3	206	7	PEM2_YEAST	METHYLENE-FATTY-ACYL-	7.26e-01
23	96	6.2	470	9	THRC_PSEAE	THREONINE SYNTHASE (E	1.29e+00
24	93	6.0	196	6	MLEF_HUMAN	MYOSIN LIGHT CHAIN 1,	3.01e+00
25	94	6.0	334	11	YCL2_ECOLI	HYPOTHETICAL 37.6 KD	2.28e+00
26	93	6.0	527	11	YAD7_SCHPO	HYPOTHETICAL 59.6 KD	3.01e+00
27	94	6.0	531	11	YP60_METTH	HYPOTHETICAL 60.5 KD	2.28e+00
28	94	6.0	534	4	FMO2_RABIT	DIMETHYLANILINE MONOO	2.28e+00
29	94	6.0	822	11	YJC9_YEAST	HYPOTHETICAL 95.4 KD	2.28e+00
30	94	6.0	885	1	AR56_SCHPO	ARG11 PROTEIN PRECURS	2.28e+00
31	94	6.0	908	6	MGR8_MOUSE	METABOTROPIC GLUTAMAT	2.28e+00
32	93	6.0	1117	11	YN96_YEAST	HYPOTHETICAL 121.1 KD	3.01e+00
33	91	5.9	2749	5	IP3R_MOUSE	INOSITOL 1,4,5-TRISPH	5.23e+00
34	91	5.9	2749	5	IP3R_RAT	INOSITOL 1,4,5-TRISPH	5.23e+00
35	90	5.8	294	6	MORA_PSEPU	MORPHINE 6-DEHYDROGEN	6.87e+00
36	90	5.8	1103	8	RPOB_CYAPA	DNA-DIRECTED RNA POLY	6.87e+00
37	89	5.7	277	7	OX40_HUMAN	OX40L RECEPTOR PRECUR	9.00e+00
38	89	5.7	388	5	HMC3_DESVH	43.2 KD PROTEIN IN HM	9.00e+00
39	89	5.7	412	11	YAG7_SCHPO	HYPOTHETICAL 45.7 KD	9.00e+00
40	89	5.7	441	1	AROA_PASMU	3-PHOSPHOSHIKIMATE 1-	9.00e+00
41	88	5.7	467	11	YG3H_YEAST	HYPOTHETICAL 54.5 KD	1.18e+01
42	89	5.7	473	10	VL2_HPV29	MINOR CAPSID PROTEIN	9.00e+00
43	89	5.7	482	9	TISD_HUMAN	TIS11D PROTEIN (EGF-R	9.00e+00
44	89	5.7	814	4	GUNE_CLOTM	ENDOGLUCANASE E PRECU	9.00e+00
45	89	5.7	912	6	MGR4_RAT	METABOTROPIC GLUTAMAT	9.00e+00

ALIGNMENTS

RESULT 1
ID BCLX_CHICK STANDARD; PRT; 190 AA.
AC Q07816;
DT 01-FEB-1995 (REL. 31, CREATED)
DT 01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)
DT 01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)
DE APOPTOSIS REGULATOR BCL-X.
GN BCL-X.
OS GALLUS GALLUS (CHICKEN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; AVES; NEOGNATHAE;
OC GALLIFORMES.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 93364977.
RA BOISE L.H., GONZALEZ-GARCIA M., POSTEMA C.E., DING L., LINDSTEN T.,
RA TURKA L.A., MAO X., NUNEZ G., THOMPSON C.B.;
RL CELL 74:597-608(1993).
CC -!- FUNCTION: DOMINANT REGULATOR OF APOPTOTIC CELL DEATH.
CC -!- TISSUE SPECIFICITY: HIGHEST EXPRESSION IN ORGANS WITH
CC LYMPHOID DEVELOPMENT.
CC -!- SIMILARITY: BELONGS TO THE BCL-2 FAMILY.
DR EMBL; Z23110; G510899; -.
DR PIR; A47537; A47537.
DR PROSITE; PS01080; BCL2.
KW APOPTOSIS.
SQ SEQUENCE 190 AA; 21467 MW; 7874E430 CRC32;

Query Match 16.6%; Score 258; DB 1; Length 190;
Best Local Similarity 28.2%; Pred. No. 1.75e-27;
Matches 37; Conservative 32; Mismatches 59; Indels 3; Gaps 3;

```
Db     60 vvngatvhrsslevheivrasdvrqalrdagdefelryrrafsdltsqlhitpgtayqsf 119
          |  |     :| ::     ::| : |   ||::: ||   |  :  :|: |:  ||: |
Qy     52 VAAPADPEMVTLPLQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENAYEYF 111

Db    120 eqvvnelfhdgvnwgrivaffsfggalcvesvdkemrvlvgrivswmttyltdh-ldpwi 178
            :   ||  | ||||:||:::||  | :   :  :  ::|::   :  ::  | :  ||
Qy    112 TKIATSLFESG-NWGRVVALLGFGYRLALHVYQHGLTGFLGQVTRFVVDFMLHHCIARWI 170

Db    179 qenggwvrtal 189
             : |||| :||
Qy    171 AQRGGWV-AAL 180
```

RESULT 2
ID BCLX_RAT STANDARD; PRT; 233 AA.

```
AC   P53563;
DT   01-OCT-1996 (REL. 34, CREATED)
DT   01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   APOPTOSIS REGULATOR BCL-X.
GN   BCLX.
OS   RATTUS NORVEGICUS (RAT).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; RODENTIA.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=BRAIN;
RA   MICHAELIDIS T.M.;
RL   SUBMITTED (DEC-1994) TO EMBL/GENBANK/DDBJ DATA BANKS.
CC   -!- FUNCTION: DOMINANT REGULATOR OF APOPTOTIC CELL DEATH.
CC   -!- ALTERNATIVE PRODUCTS: TWO ISOFORMS, BCX-X(L) (SHOWN HERE) AND
CC       BCL-X(S), ARE DERIVED BY ALTERNATIVE SPLICING.
CC   -!- SIMILARITY: BELONGS TO THE BCL-2 FAMILY.
DR   EMBL; X82537; G607177; -.
DR   EMBL; X82537; G607178; -.
KW   APOPTOSIS; ALTERNATIVE SPLICING.
FT   VARSPLIC    126    188       MISSING (IN BCL-X(S)).
SQ   SEQUENCE   233 AA;  26130 MW;  E0589815 CRC32;

  Query Match           16.3%;  Score 254;  DB 1;  Length 233;
  Best Local Similarity  25.5%;  Pred. No. 9.85e-27;
  Matches    36;  Conservative   41;  Mismatches  60;  Indels   4;  Gaps   4;

Db     61 dspavngatghs-ssldarevipmaavkqalreagdefelryrrafsdltsqlhitpgta 119
          :: :| :::     :|  :    |: | : |   ||::: ||   |  :  :|: |:  |
Qy     48 EAEGVAAPADPEMVTLPLQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENA 107

Db    120 yqsfeqvvnelfrdgvnwgrivaffsfggalcvesvdkemqvlvsriaswmatylndh-l 178
          |: |  :   ||  | ||||:||:::||  | :   :  :  ::::::  :  ::  | :
Qy    108 YEYFTKIATSLFESG-NWGRVVALLGFGYRLALHVYQHGLTGFLGQVTRFVVDFMLHHCI 166

Db    179 epwiqenggwdtfvdlygnna 199
            || : ||| : ::| || :
Qy    167 ARWIAQRGGWVAALNL-GNGP 186


RESULT    3
ID   BCLX_HUMAN     STANDARD;      PRT;   233 AA.
AC   Q07817;
DT   01-FEB-1995 (REL. 31, CREATED)
DT   01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   APOPTOSIS REGULATOR BCL-X.
GN   BCLX.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 93364977.
RA   BOISE L.H., GONZALEZ-GARCIA M., POSTEMA C.E., DING L.,
RA   LINDSTEN T., TURKA L.A., MAO X., NUNEZ G., THOMPSON C.B.;
RL   CELL 74:597-608(1993).
CC   -!- FUNCTION: DOMINANT REGULATOR OF APOPTOTIC CELL DEATH.
CC   -!- ALTERNATIVE PRODUCTS: TWO ISOFORMS, BCX-X(L) (SHOWN HERE) AND
CC       BCL-X(S), ARE DERIVED BY ALTERNATIVE SPLICING.
CC   -!- TISSUE SPECIFICITY: BCL-X(S) IS EXPRESSED AT HIGH LEVELS IN CELLS
CC       THAT UNDERGO A HIGH RATE OF TURNOVER, SUCH AS DEVELOPING
CC       LYMPHOCYTES. IN CONTRAST, BCL-X(L) IS FOUND IN TISSUES CONTAINING
CC       LONG-LIVED POSTMITOTIC CELLS, SUCH AS ADULT BRAIN.
CC   -!- SIMILARITY: BELONGS TO THE BCL-2 FAMILY.
DR   EMBL; Z23116; G623237; -.
DR   EMBL; Z23115; G510901; -.
DR   MIM; 600039; -.
DR   PROSITE; PS01080; BCL2.
KW   APOPTOSIS; ALTERNATIVE SPLICING.
FT   VARSPLIC    126    188       MISSING (IN BCL-X(S)).
FT   CONFLICT     70     70       G -> A (IN G510901).
SQ   SEQUENCE   233 AA;  26049 MW;  57C67491 CRC32;

  Query Match           16.1%;  Score 250;  DB 1;  Length 233;
  Best Local Similarity  25.5%;  Pred. No. 5.51e-26;
  Matches    36;  Conservative   41;  Mismatches  60;  Indels   4;  Gaps   4;

Db     61 dspavngatghs-ssldarevipmaavkqalreagdefelryrrafsdltsqlhitpgta 119
          :: :| :::     :|  :    |: | : |   ||::: ||   |  :  :|: |:  |
Qy     48 EAEGVAAPADPEMVTLPLQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENA 107

Db    120 yqsfeqvvnelfrdgvnwgrivaffsfggalcvesvdkemqvlvsriaawmatylndh-l 178
          |: |  :   ||  | ||||:||:::||  | :   :  :  ::::::  :  ::  | :
Qy    108 YEYFTKIATSLFESG-NWGRVVALLGFGYRLALHVYQHGLTGFLGQVTRFVVDFMLHHCI 166

Db    179 epwiqenggwdtfvelygnna 199
            || : ||| : ::| || :
Qy    167 ARWIAQRGGWVAALNL-GNGP 186


RESULT    4
ID   BCL2_CHICK     STANDARD;      PRT;   233 AA.
AC   Q00709;
DT   01-APR-1993 (REL. 25, CREATED)
DT   01-APR-1993 (REL. 25, LAST SEQUENCE UPDATE)
DT   01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE   APOPTOSIS REGULATOR BCL-2.
GN   BCL-2.
OS   GALLUS GALLUS (CHICKEN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; AVES; NEOGNATHAE;
OC   GALLIFORMES.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 92375724.
RA   EGUCHI Y., EWERT D.L., TSUJIMOTO Y.;
RL   NUCLEIC ACIDS RES. 20:4187-4192(1992).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 92379084.
RA   CAZALS-HATEM D.L., LOUIE D.C., TANAKA S., REED J.C.;
RL   BIOCHIM. BIOPHYS. ACTA 1132:109-113(1992).
CC   -!- FUNCTION: PROLONGS THE SURVIVAL OF HEMATOPOIETIC CELLS IN THE
CC       ABSENCE OF REQUIRED GROWTH FACTORS AND ALSO IN THE PRESENCE OF
CC       VARIOUS STIMULI INDUCING CELLULAR DEATH. BLOCKS APOPTOSIS. MIGHT
CC       FUNCTION IN AN ANTIOXIDANT PATHWAY TO PREVENT APOPTOSIS AT SITES
CC       OF FREE RADICAL GENERATION SUCH AS MITOCHONDRIA.
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL INNER MEMBRANE
CC       (BY SIMILARITY).
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN A VARIETY OF TISSUES.
CC   -!- SIMILARITY: BELONGS TO THE BCL-2 FAMILY.
DR   EMBL; D11382; G222794; -.
DR   EMBL; D11381; G222794; JOINED.
DR   EMBL; Z11961; G62970; -.
DR   PIR; A37332; A37332.
DR   PIR; S24390; S24390.
DR   PROSITE; PS01080; BCL2.
KW   APOPTOSIS; TRANSMEMBRANE; MITOCHONDRION.
FT   TRANSMEM    208    228       POTENTIAL.
FT   CONFLICT     64     64       E -> S (IN REF. 2).
FT   CONFLICT     67     82       GSAAASEVPPAEGLRP -> ARLLLVRCPRLRGCA
FT                                (IN REF. 2).
FT   CONFLICT    121    121       H -> T (IN REF. 2).
FT   CONFLICT    139    139       G -> V (IN REF. 2).
SQ   SEQUENCE   233 AA;  25687 MW;  3376502C CRC32;

  Query Match           15.8%;  Score 245;  DB 1;  Length 233;
  Best Local Similarity  28.5%;  Pred. No. 4.70e-25;
  Matches    41;  Conservative   40;  Mismatches  57;  Indels   6;  Gaps   6;

Db     60 hhrpeppgsaaasevppae-glrpapp-g-vhlalrqagdefsrryqrdfaqmsgqlhlt 116
          ::  |: | ||:::       |:|::  | |   |   ||:::|||: :|  |  :|: |
Qy     44 QQEQEAEGVAAPADPEMVTLPLQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPT 103

Db    117 pftahgrfvavveelfrdgvnwgrivaffefggvmcvesvnremsplvdniatwmteyln 176
```

```
          :  |   |  :   ||  | ||||:||:: ||  : :    : :: ::  ::  : :::
Qy    104 AENAYEYFTKIATSLFESG-NWGRVVALLGFGYRLALHVYQHGLTGFLGQVTRFVVDFML 162

Db    177 rh-lhnwiqdnggwdafvelygns 199
          :| :  || : ||| | ::| ||:
Qy    163 HHCIARWIAQRGGWVAALNL-GNG 185


RESULT    5
ID   BC2B_HUMAN     STANDARD;      PRT;   205 AA.
AC   P10416;
DT   01-MAR-1989 (REL. 10, CREATED)
DT   01-APR-1993 (REL. 25, LAST SEQUENCE UPDATE)
DT   01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE   PROTEIN BCL-2-BETA.
GN   BCL2.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 86259760.
RA   TSUJIMOTO Y., CROCE C.M.;
RL   PROC. NATL. ACAD. SCI. U.S.A. 83:5214-5218(1986).
RN   [2]
RP   REVISIONS TO 96 AND 110.
RX   MEDLINE; 92375724.
RA   EGUCHI Y., EWERT D.L., TSUJIMOTO Y.;
RL   NUCLEIC ACIDS RES. 20:4187-4192(1992).
RN   [3]
RP   SUBCELLULAR LOCATION.
RX   MEDLINE; 91066924.
RA   HOCKENBERY D., NUNEZ G., MILLIMAN C., SCHREIBER R.D., KORSMEYER S.J.;
RL   NATURE 348:334-336(1990).
CC   -!- FUNCTION: PROLONGS THE SURVIVAL OF HEMATOPOIETIC CELLS IN THE
CC       ABSENCE OF REQUIRED GROWTH FACTORS AND ALSO IN THE PRESENCE OF
CC       VARIOUS STIMULI INDUCING CELLULAR DEATH. BLOCKS APOPTOSIS. MIGHT
CC       FUNCTION IN AN ANTIOXIDANT PATHWAY TO PREVENT APOPTOSIS AT SITES
CC       OF FREE RADICAL GENERATION SUCH AS MITOCHONDRIA.
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL INNER MEMBRANE.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN A VARIETY OF TISSUES.
CC   -!- DISEASE: INVOLVED IN FOLLICULAR LYMPHOMA (FL) (ALSO KNOWN AS TYPE
CC       II CHRONIC LYMPHATIC LEUKEMIA) BY A CHROMOSOMAL TRANSLOCATION
CC       T(14;18)(Q32;Q21) WHICH INVOLVES BCL2 AND IMMUNOGLOBULIN GENE
CC       REGIONS.
CC   -!- ALTERNATIVE PRODUCTS: TWO FORMS OF BCL-2: ALPHA, AND BETA, ARE
CC       PRODUCED BY ALTERNATIVE SPLICING OF THE SAME GENE. THEY ONLY
CC       DIFFER AT THEIR C-TERMINAL ENDS.
CC   -!- SIMILARITY: BELONGS TO THE BCL-2 FAMILY.
DR   EMBL; M13995; G179369; ALT_SEQ.
DR   PIR; B29409; TVHUB1.
DR   PIR; D37332; D37332.
DR   MIM; 151430; -.
DR   PROSITE; PS01080; BCL2.
KW   PROTO-ONCOGENE; APOPTOSIS; ALTERNATIVE SPLICING; MEMBRANE;
KW   MITOCHONDRION; CHROMOSOMAL TRANSLOCATION.
SQ   SEQUENCE   205 AA;  22311 MW;  ED321E5E CRC32;

  Query Match           15.4%;  Score 239;  DB 1;  Length 205;
  Best Local Similarity  29.7%;  Pred. No. 6.08e-24;
  Matches    38;  Conservative   34;  Mismatches  51;  Indels   5;  Gaps   4;

Db     74 tpaapgaaagpalspvppvvhlt--lrqagddfsrryrrdfaemssqlhltpftargrfa 131
          :|| |   : | | | : : ::   |   |||::|||  :|  |  :|: |:  |  |:
Qy     54 APADPEMVTLP-LQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENAYEYFT 112

Db    132 tvveelfrdgvnwgrivaffefggvmcvesvnremsplvdnialwmteylnrh-lhtwiq 190
          :   ||  | ||||:||:: ||  : :    : :: ::  ::  : ::: :| :  ||
Qy    113 KIATSLFESG-NWGRVVALLGFGYRLALHVYQHGLTGFLGQVTRFVVDFMLHHCIARWIA 171

Db    191 dnggwvga 198
          : ||||:|
Qy    172 QRGGWVAA 179
```

```
RESULT    6
ID   BC2A_HUMAN     STANDARD;      PRT;   239 AA.
AC   P10415;
DT   01-MAR-1989 (REL. 10, CREATED)
DT   01-APR-1993 (REL. 25, LAST SEQUENCE UPDATE)
DT   01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE   PROTEIN BCL-2-ALPHA.
GN   BCL2.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 86259760.
RA   TSUJIMOTO Y., CROCE C.M.;
RL   PROC. NATL. ACAD. SCI. U.S.A. 83:5214-5218(1986).
RN   [2]
RP   REVISIONS TO 96; 110 AND 237.
RX   MEDLINE; 92375724.
RA   EGUCHI Y., EWERT D.L., TSUJIMOTO Y.;
RL   NUCLEIC ACIDS RES. 20:4187-4192(1992).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 87002488.
RA   CLEARY M.L., SMITH S.D., SKLAR J.;
RL   CELL 47:19-28(1986).
RN   [4]
RP   SUBCELLULAR LOCATION.
RX   MEDLINE; 91066924.
RA   HOCKENBERY D., NUNEZ G., MILLIMAN C., SCHREIBER R.D., KORSMEYER S.J.;
RL   NATURE 348:334-336(1990).
CC   -!- FUNCTION: PROLONGS THE SURVIVAL OF HEMATOPOIETIC CELLS IN THE
CC       ABSENCE OF REQUIRED GROWTH FACTORS AND ALSO IN THE PRESENCE OF
CC       VARIOUS STIMULI INDUCING CELLULAR DEATH. BLOCKS APOPTOSIS. MIGHT
CC       FUNCTION IN AN ANTIOXIDANT PATHWAY TO PREVENT APOPTOSIS AT SITES
CC       OF FREE RADICAL GENERATION SUCH AS MITOCHONDRIA.
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL INNER MEMBRANE.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN A VARIETY OF TISSUES.
CC   -!- DISEASE: INVOLVED IN FOLLICULAR LYMPHOMA (FL) (ALSO KNOWN AS TYPE
CC       II CHRONIC LYMPHATIC LEUKEMIA) BY A CHROMOSOMAL TRANSLOCATION
CC       T(14;18)(Q32;Q21) WHICH INVOLVES BCL2 AND IMMUNOGLOBULIN GENE
CC       REGIONS.
CC   -!- ALTERNATIVE PRODUCTS: TWO FORMS OF BCL-2: ALPHA, AND BETA, ARE
CC       PRODUCED BY ALTERNATIVE SPLICING OF THE SAME GENE. THEY ONLY
CC       DIFFER AT THEIR C-TERMINAL ENDS.
CC   -!- SIMILARITY: BELONGS TO THE BCL-2 FAMILY.
DR   EMBL; M13994; G179367; ALT_SEQ.
DR   EMBL; M14745; G179371; -.
DR   PIR; A29409; TVHUA1.
DR   PIR; A24428; TVHUBC.
DR   PIR; C37332; C37332.
DR   MIM; 151430; -.
DR   PROSITE; PS01080; BCL2.
KW   PROTO-ONCOGENE; APOPTOSIS; ALTERNATIVE SPLICING; TRANSMEMBRANE;
KW   MITOCHONDRION; CHROMOSOMAL TRANSLOCATION.
FT   TRANSMEM    212    233       POTENTIAL.
FT   CONFLICT     59     59       P -> T (IN REF. 3).
FT   CONFLICT    117    117       S -> R (IN REF. 3).
SQ   SEQUENCE   239 AA;  26266 MW;  75084B59 CRC32;

  Query Match           15.4%;  Score 239;  DB 1;  Length 239;
  Best Local Similarity  29.3%;  Pred. No. 6.08e-24;
  Matches    41;  Conservative   36;  Mismatches  57;  Indels   6;  Gaps   5;

Db     74 tpaapgaaagpalspvppvvhlt--lrqagddfsrryrrdfaemssqlhltpftargrfa 131
          :|| |   : | | | : : ::   |   |||::|||  :|  |  :|: |:  |  |:
Qy     54 APADPEMVTLP-LQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAENAYEYFT 112

Db    132 tvveelfrdgvnwgrivaffefggvmcvesvnremsplvdnialwmteylnrh-lhtwiq 190
          :   ||  | ||||:||:: ||  : :    : :: ::  ::  : ::: :| :  ||
Qy    113 KIATSLFESG-NWGRVVALLGFGYRLALHVYQHGLTGFLGQVTRFVVDFMLHHCIARWIA 171
```

```
Db    191 dnggwdafvel-ygpsmrpl 209
          : ||| | ::|  || :  |
Qy    172 QRGGWVAALNLGNGPILNVL 191


RESULT    7
ID   BC2B_MOUSE     STANDARD;      PRT;   199 AA.
AC   P10418;
DT   01-MAR-1989 (REL. 10, CREATED)
DT   01-MAR-1989 (REL. 10, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   PROTEIN BCL-2-BETA.
GN   BCL2 OR BCL-2.
OS   MUS MUSCULUS (MOUSE).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; RODENTIA.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 87187643.
RA   NEGRINI M., SILINI E., KOZAK C., TSUJIMOTO Y., CROCE C.M.;
RL   CELL 49:455-463(1987).
CC   -!- FUNCTION: PROLONGS THE SURVIVAL OF HEMATOPOIETIC CELLS IN THE
CC       ABSENCE OF REQUIRED GROWTH FACTORS AND ALSO IN THE PRESENCE OF
CC       VARIOUS STIMULI INDUCING CELLULAR DEATH. BLOCKS APOPTOSIS. MIGHT
CC       FUNCTION IN AN ANTIOXIDANT PATHWAY TO PREVENT APOPTOSIS AT SITES
CC       OF FREE RADICAL GENERATION SUCH AS MITOCHONDRIA.
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL INNER MEMBRANE.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN A VARIETY OF TISSUES.
CC   -!- ALTERNATIVE PRODUCTS: TWO FORMS OF BCL-2: ALPHA, AND BETA, ARE
CC       PRODUCED BY ALTERNATIVE SPLICING OF THE SAME GENE. THEY ONLY
CC       DIFFER AT THEIR C-TERMINAL ENDS.
CC   -!- SIMILARITY: BELONGS TO THE BCL-2 FAMILY.
DR   EMBL; M16506; G387110; -.
DR   PIR; B25960; TVMSB1.
DR   PROSITE; PS01080; BCL2.
KW   APOPTOSIS; ALTERNATIVE SPLICING; MEMBRANE; MITOCHONDRION.
SQ   SEQUENCE   199 AA;  22299 MW;  3E60FA47 CRC32;

  Query Match           15.1%;  Score 234;  DB 1;  Length 199;
  Best Local Similarity  31.1%;  Pred. No. 5.08e-23;
  Matches    32;  Conservative   28;  Mismatches  41;  Indels   2;  Gaps   2;

Db     94 lrragddfsrryrrdfaemssqlhltpftargrfatvveelfrdgvnwgrivaffefggv 153
          |   |||::|||  :|  |  :|: |:  |   |: :   ||  | ||||:||:: ||
Qy     78 LAIIGDDINRRYDSEFQTMLQHLQPTAENAYEYFTKIATSLFESG-NWGRVVALLGFGYR 136

Db    154 mcvesvnremsplvdnialwmteylnrh-lhtwiqdnggwvga 195
          : :    : :: ::  ::  : ::: :| :  || : ||||:|
Qy    137 LALHVYQHGLTGFLGQVTRFVVDFMLHHCIARWIAQRGGWVAA 179


RESULT    8
ID   BC2A_MOUSE     STANDARD;      PRT;   236 AA.
AC   P10417;
DT   01-MAR-1989 (REL. 10, CREATED)
DT   01-APR-1993 (REL. 25, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   PROTEIN BCL-2 ALPHA.
GN   BCL2 OR BCL-2.
OS   MUS MUSCULUS (MOUSE).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; RODENTIA.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/C; TISSUE=LIVER;
RX   MEDLINE; 87187643.
RA   NEGRINI M., SILINI E., KOZAK C., TSUJIMOTO Y., CROCE C.M.;
RL   CELL 49:455-463(1987).
RN   [2]
RP   REVISIONS TO 221-222.
RX   MEDLINE; 92375724.
RA   EGUCHI Y., EWERT D.L., TSUJIMOTO Y.;
RL   NUCLEIC ACIDS RES. 20:4187-4192(1992).
CC   -!- FUNCTION: PROLONGS THE SURVIVAL OF HEMATOPOIETIC CELLS IN THE
CC       ABSENCE OF REQUIRED GROWTH FACTORS AND ALSO IN THE PRESENCE OF
CC       VARIOUS STIMULI INDUCING CELLULAR DEATH. BLOCKS APOPTOSIS. MIGHT
CC       FUNCTION IN AN ANTIOXIDANT PATHWAY TO PREVENT APOPTOSIS AT SITES
CC       OF FREE RADICAL GENERATION SUCH AS MITOCHONDRIA.
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL INNER MEMBRANE.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN A VARIETY OF TISSUES.
CC   -!- ALTERNATIVE PRODUCTS: TWO FORMS OF BCL-2: ALPHA, AND BETA, ARE
CC       PRODUCED BY ALTERNATIVE SPLICING OF THE SAME GENE. THEY ONLY
CC       DIFFER AT THEIR C-TERMINAL ENDS.
CC   -!- SIMILARITY: BELONGS TO THE BCL-2 FAMILY.
DR   EMBL; L31532; G387109; -.
DR   EMBL; M16506; G387109; JOINED.
DR   PIR; A25960; TVMSA1.
DR   PIR; E37332; E37332.
DR   PROSITE; PS01080; BCL2.
KW   APOPTOSIS; ALTERNATIVE SPLICING; TRANSMEMBRANE; MITOCHONDRION.
FT   TRANSMEM    209    230       POTENTIAL.
SQ   SEQUENCE   236 AA;  26425 MW;  7ADFE975 CRC32;

  Query Match           15.1%;  Score 234;  DB 1;  Length 236;
  Best Local Similarity  30.4%;  Pred. No. 5.08e-23;
  Matches    35;  Conservative   30;  Mismatches  47;  Indels   3;  Gaps   3;

Db     94 lrragddfsrryrrdfaemssqlhltpftargrfatvveelfrdgvnwgrivaffefggv 153
          |   |||::|||  :|  |  :|: |:  |   |: :   ||  | ||||:||:: ||
Qy     78 LAIIGDDINRRYDSEFQTMLQHLQPTAENAYEYFTKIATSLFESG-NWGRVVALLGFGYR 136

Db    154 mcvesvnremsplvdnialwmteylnrh-lhtwiqdnggwdafvel-ygpsmrpl 206
          : :    : :: ::  ::  : ::: :| :  || : ||| | ::|  || :  |
Qy    137 LALHVYQHGLTGFLGQVTRFVVDFMLHHCIARWIAQRGGWVAALNLGNGPILNVL 191


RESULT    9
ID   BCL2_RAT       STANDARD;      PRT;   236 AA.
AC   P49950;
DT   01-OCT-1996 (REL. 34, CREATED)
DT   01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   PROTEIN BCL-2 ALPHA.
GN   BCL2 OR BCL-2.
OS   RATTUS NORVEGICUS (RAT).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; RODENTIA.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=BRAIN;
RX   MEDLINE; 94193015.
RA   SATO T., IRIE S., KRAJEWSKI S., REED J.C.;
RL   GENE 140:291-292(1994).
CC   -!- FUNCTION: PROLONGS THE SURVIVAL OF HEMATOPOIETIC CELLS IN THE
CC       ABSENCE OF REQUIRED GROWTH FACTORS AND ALSO IN THE PRESENCE OF
CC       VARIOUS STIMULI INDUCING CELLULAR DEATH. BLOCKS APOPTOSIS. MIGHT
CC       FUNCTION IN AN ANTIOXIDANT PATHWAY TO PREVENT APOPTOSIS AT SITES
CC       OF FREE RADICAL GENERATION SUCH AS MITOCHONDRIA.
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL INNER MEMBRANE.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN A VARIETY OF TISSUES (BY
CC       SIMILARITY).
CC   -!- ALTERNATIVE PRODUCTS: TWO FORMS OF BCL-2: ALPHA, AND BETA, ARE
CC       PRODUCED BY ALTERNATIVE SPLICING OF THE SAME GENE. THEY ONLY
CC       DIFFER AT THEIR C-TERMINAL ENDS.
CC   -!- SIMILARITY: BELONGS TO THE BCL-2 FAMILY.
DR   EMBL; L14680; G408947; -.
KW   APOPTOSIS; ALTERNATIVE SPLICING; TRANSMEMBRANE; MITOCHONDRION.
FT   TRANSMEM    209    230       POTENTIAL.
SQ   SEQUENCE   236 AA;  26550 MW;  336E6B40 CRC32;

  Query Match           14.9%;  Score 232;  DB 1;  Length 236;
  Best Local Similarity  27.2%;  Pred. No. 1.18e-22;
  Matches    40;  Conservative   42;  Mismatches  60;  Indels   5;  Gaps   5;

Db     63 rdtaartsplrp-lvanagpalspvppvvhltlrragddfsrryrrdfaemssqlhltpf 121
```

```
          ::: : ::|  | :|: :  : | :  | :  |  |||::|||  :|  |  :|: |:
Qy     47 QEAEGVAAPADPEMVTLPLQPSSTMGQVGR-QLAIIGDDINRRYDSEFQTMLQHLQPTAE 105

Db    122 targrfatvveelfrdgvnwgrivaffefggvmcvgsvnremsplvdnialwmteylnrh 181
          |   |: :   ||  | ||||:||:: ||  : :    : :: ::  ::  : ::: :|
Qy    106 NAYEYFTKIATSLFESG-NWGRVVALLGFGYRLALHVYQHGLTGFLGQVTRFVVDFMLHH 164

Db    182 -lhtwiqdnggwdafvel-ygpsmrpl 206
           :  || : ||| | ::|  || :  |
Qy    165 CIARWIAQRGGWVAALNLGNGPILNVL 191


RESULT   10
ID   BAXA_MOUSE     STANDARD;      PRT;   192 AA.
AC   Q07813;
DT   01-FEB-1995 (REL. 31, CREATED)
DT   01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   APOPTOSIS REGULATOR BAX, MEMBRANE ISOFORM ALPHA.
GN   BAX.
OS   MUS MUSCULUS (MOUSE).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; RODENTIA.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6 X DBA/2 F1;
RX   MEDLINE; 93364978.
RA   OLTVAI Z.N., MILLIMAN C.L., KORSMEYER S.J.;
RL   CELL 74:609-619(1993).
CC   -!- FUNCTION: ACCELERATES PROGRAMMED CELL DEATH BY BINDING TO THE
CC       APOPTOSIS REPRESSOR BCL-2.
CC   -!- ALTERNATIVE PRODUCTS: A 21 KD MEMBRANE PROTEIN ALPHA AND THE TWO
CC       CYTOPLASMIC PROTEINS BETA AND GAMMA ARE GENERATED BY ALTERNATIVE
CC       SPLICING.
CC   -!- SUBUNIT: FORMS HOMODIMERS AND HETERODIMERS TOGETHER WITH BCL-2.
CC   -!- SUBCELLULAR LOCATION: MEMBRANE.
CC   -!- SIMILARITY: BELONGS TO THE BCL-2 FAMILY.
DR   EMBL; L22472; G388192; -.
DR   PROSITE; PS01080; BCL2.
KW   APOPTOSIS; TRANSMEMBRANE; ALTERNATIVE SPLICING.
FT   TRANSMEM    172    192       POTENTIAL.
SQ   SEQUENCE   192 AA;  21394 MW;  BD035304 CRC32;

  Query Match           13.3%;  Score 206;  DB 1;  Length 192;
  Best Local Similarity  25.6%;  Pred. No. 5.98e-18;
  Matches    46;  Conservative   49;  Mismatches  75;  Indels  10;  Gaps   8;

Db     16 seqimktgafllqgfiqdragrmagetpeltleqppqdastk-klseclrrigdelds-- 72
          :|:: ::  |  :   |:  | |   ||: :  | | :||:  ::  |  |||:::
Qy     31 TEEVFRSYVFYRHQQEQEAEGVAAPADPEM-VTLPLQPSSTMGQVGRQLAIIGDDINRRY 89

Db     73 nmelqrmiadvdtdspre-vff-rvaadmfadgnfnwgrvvalfyfasklvlkalctkvp 130
          : |:| |:  ::  :     :| ::|: :|  ||  |||||||: |: :| |       :
Qy     90 DSEFQTMLQHLQPTAENAYEYFTKIATSLFESGN--WGRVVALLGFGYRLALHVYQHGLT 147

Db    131 elirtimgwtldf-lrerllvwiqdqggwegllsyfgtptwqtvtifvagvltasltiwk 189
           ::  :    :|| |:  :  || ::||| : |:    |    | : :: || : : :::
Qy    148 GFLGQVTRFVVDFMLHHCIARWIAQRGGWVAALNLGNGPI-LNVLVVLGVVLLGQFVVRR 206


RESULT   11
ID   BAXA_HUMAN     STANDARD;      PRT;   192 AA.
AC   Q07812;
DT   01-FEB-1995 (REL. 31, CREATED)
DT   01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   APOPTOSIS REGULATOR BAX, MEMBRANE ISOFORM ALPHA.
GN   BAX.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=B-CELL;
RX   MEDLINE; 93364978.
RA   OLTVAI Z.N., MILLIMAN C.L., KORSMEYER S.J.;
RL   CELL 74:609-619(1993).
CC   -!- FUNCTION: ACCELERATES PROGRAMMED CELL DEATH BY BINDING TO THE
CC       APOPTOSIS REPRESSOR BCL-2.
CC   -!- SUBUNIT: FORMS HOMODIMERS AND HETERODIMERS TOGETHER WITH BCL-2.
CC   -!- SUBCELLULAR LOCATION: MEMBRANE.
CC   -!- ALTERNATIVE PRODUCTS: THE MEMBRANE ISOFORM ALPHA AND THE THREE
CC       CYTOPLASMIC ISOFORMS, BETA, GAMMA AND DELTA ARE GENERATED BY
CC       ALTERNATIVE SPLICING.
CC   -!- SIMILARITY: BELONGS TO THE BCL-2 FAMILY.
DR   EMBL; L22473; G388166; -.
DR   PIR; A47538; A47538.
DR   MIM; 600040; -.
DR   PROSITE; PS01080; BCL2.
KW   APOPTOSIS; TRANSMEMBRANE; ALTERNATIVE SPLICING.
FT   TRANSMEM    172    192       POTENTIAL.
SQ   SEQUENCE   192 AA;  21184 MW;  B2E6148A CRC32;

  Query Match           12.5%;  Score 194;  DB 1;  Length 192;
  Best Local Similarity  24.5%;  Pred. No. 7.87e-16;
  Matches    40;  Conservative   44;  Mismatches  71;  Indels   8;  Gaps   6;

Db     32 qdragrmggeapelaldpvpqdastkklseclkrigdelds--nmelqrmiaavdtdspr 89
          |: |  :  ||:   |:   ::   ::  |  |||:::   : |:| |:  ::  :
Qy     47 QEAEGVAAPADPEMVTLPLQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAEN 106

Db     90 e-vff-rvaadmfsdgnfnwgrvvalfyfasklvlkalctkvpelirtimgwtldf-lre 146
             :| ::|: :|  ||  |||||||: |: :| |       :  ::  :    :|| |:
Qy    107 AYEYFTKIATSLFESGN--WGRVVALLGFGYRLALHVYQHGLTGFLGQVTRFVVDFMLHH 164

Db    147 rllgwiqdqggwdgllsyfgtptwqtvtifvagvltasltiwk 189
           :  || ::||| : |:    |    | : :: || : : :::
Qy    165 CIARWIAQRGGWVAALNLGNGPI-LNVLVVLGVVLLGQFVVRR 206


RESULT   12
ID   BAXB_HUMAN     STANDARD;      PRT;   218 AA.
AC   Q07814;
DT   01-FEB-1995 (REL. 31, CREATED)
DT   01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   APOPTOSIS REGULATOR BAX, CYTOPLASMIC ISOFORM BETA.
GN   BAX.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=B-CELL;
RX   MEDLINE; 93364978.
RA   OLTVAI Z.N., MILLIMAN C.L., KORSMEYER S.J.;
RL   CELL 74:609-619(1993).
CC   -!- FUNCTION: ACCELERATES PROGRAMMED CELL DEATH BY BINDING TO THE
CC       APOPTOSIS REPRESSOR BCL-2.
CC   -!- SUBUNIT: FORMS HOMODIMERS AND HETERODIMERS TOGETHER WITH BCL-2.
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC   -!- ALTERNATIVE PRODUCTS: THE MEMBRANE ISOFORM ALPHA AND THE THREE
CC       CYTOPLASMIC ISOFORMS, BETA, GAMMA AND DELTA ARE GENERATED BY
CC       ALTERNATIVE SPLICING.
CC   -!- SIMILARITY: BELONGS TO THE BCL-2 FAMILY.
DR   EMBL; L22474; G388168; -.
DR   PIR; B47538; B47538.
DR   MIM; 600040; -.
DR   PROSITE; PS01080; BCL2.
KW   APOPTOSIS; ALTERNATIVE SPLICING.
SQ   SEQUENCE   218 AA;  24220 MW;  82B2FF09 CRC32;

  Query Match           12.2%;  Score 189;  DB 1;  Length 218;
  Best Local Similarity  27.0%;  Pred. No. 5.86e-15;
  Matches    37;  Conservative   35;  Mismatches  58;  Indels   7;  Gaps   5;
```

```
Db      32 qdragrmggeapelaldpvpqdastkklseclkrigdelds--nmelqrmiaavdtdspr 89
           |: |   :    ||:   |:   ::   ::  |  |||:::   : |:| |:  ::  :
Qy      47 QEAEGVAAPADPEMVTLPLQPSSTMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQPTAEN 106

Db      90 e-vff-rvaadmfsdgnfnwgrvvalfyfasklvlkalctkvpelirtimgwtldf-lre 146
             :| ::|: :|  ||  |||||||: |: :| |      :  ::  :    :|| |:
Qy     107 AYEYFTKIATSLFESGN--WGRVVALLGFGYRLALHVYQHGLTGFLGQVTRFVVDFMLHH 164

Db     147 rllgwiqdqggwvrllk 163
           :  || ::|||| |:
Qy     165 CIARWIAQRGGWVAALN 181


RESULT   13
ID   MCL1_HUMAN     STANDARD;      PRT;   350 AA.
AC   Q07820;
DT   01-FEB-1995 (REL. 31, CREATED)
DT   01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)
DT   01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)
DE   INDUCED MYELOID LEUKEMIA CELL DIFFERENTIATION PROTEIN MCL1.
GN   MCL1.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=MYELOID LEUKEMIA CELLS;
RX   MEDLINE; 93234528.
RA   KOZOPAS K.M., YANG T., BUCHAN H.L., ZHOU P., CRAIG R.W.;
RL   PROC. NATL. ACAD. SCI. U.S.A. 90:3516-3520(1993).
CC   -!- FUNCTION: INVOLVED IN PROGRAMMING OF DIFFERENTIATION AND
CC       CONCOMITANT MAINTENANCE OF VIABILITY BUT NOT OF PROLIFERATION
CC       (PROBABLE).
CC   -!- INDUCTION: EXPRESSION INCREASES EARLY DURING PHORBOL-ESTER INDUCED
CC       DIFFERENTIATION ALONG THE MONOCYTE/MAGROPHAGE PATHWAY IN MYELOID
CC       LEUKEMIA CELL LINES ML-1.
CC   -!- SIMILARITY: BELONGS TO THE BCL2 FAMILY.
DR   EMBL; L08246; -; NOT_ANNOTATED_CDS.
DR   PIR; A47476; A47476.
DR   MIM; 159552; -.
DR   PROSITE; PS01080; BCL2.
KW   APOPTOSIS; TRANSMEMBRANE; DIFFERENTIATION.
FT   UNSURE      227    227       OR A.
FT   TRANSMEM    330    349       POTENTIAL.
SQ   SEQUENCE   350 AA;  37365 MW;  10194B64 CRC32;

  Query Match           11.3%;  Score 176;  DB 6;  Length 350;
  Best Local Similarity  24.3%;  Pred. No. 1.00e-12;
  Matches   34;  Conservative   39;  Mismatches  63;  Indels   4;  Gaps   3;

Db     174 lyrqsleiisrylreqatgakdtkpmgrsgatsrkaletlrrvgdgvqrnhetvfqgmlr 233
           :||:  |  :  :   |   : | |:  |:: :: : : |  :|| : |  :: || ||:
Qy      40 FYRHQQEQEAEGVAAPADPEMVTLPLQPSSTMGQVGRQ-LAIIGDDINRRYDSEFQTMLQ 98

Db     234 kldikneddvkslsrvmihvfsdgvtnwgrivtlisfgafvakhlktinqescieplaes 293
             |:   |:    ::::   :|  |  ||||:|:|::||  :| |:      : :  ::
Qy      99 HLQPTAENAYEYFTKIATSLFESG--NWGRVVALLGFGYRLALHVYQHGLTGFLGQVTRF 156

Db     294 itd-vlvrtkrdwlvkqrgw 312
           : | :| :    |:  : ||
Qy     157 VVDFMLHHCIARWIAQRGGW 176


RESULT   14
ID   BAXD_HUMAN     STANDARD;      PRT;   143 AA.
AC   P55269;
DT   01-OCT-1996 (REL. 34, CREATED)
DT   01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   BAX PROTEIN, CYTOPLASMIC ISOFORM DELTA.
GN   BAX.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 95331797.
RA   APTE S.S., MATTEI M.-G., OLSEN B.R.;
RL   GENOMICS 26:592-594(1995).
CC   -!- ALTERNATIVE PRODUCTS: THE MEMBRANE ISOFORM ALPHA AND THE THREE
CC       CYTOPLASMIC ISOFORMS, BETA, GAMMA AND DELTA ARE GENERATED BY
CC       ALTERNATIVE SPLICING.
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC (POTENTIAL).
CC   -!- SIMILARITY: BELONGS TO THE BCL-2 FAMILY.
DR   EMBL; U19599; G841238; -.
DR   MIM; 600040; -.
KW   APOPTOSIS; ALTERNATIVE SPLICING.
SQ   SEQUENCE   143 AA;  15772 MW;  71AA1CBD CRC32;

  Query Match           10.6%;  Score 165;  DB 1;  Length 143;
  Best Local Similarity  27.4%;  Pred. No. 7.08e-11;
  Matches   29;  Conservative   30;  Mismatches  43;  Indels   4;  Gaps   3;

Db      36 tdsprevffrvaadmfsdgnfnwgrvvalfyfasklvlkalctkvpelirtimgwtldf- 94
           :::: | | ::|: :|  ||  |||||||: |: :| |      :  ::  :    :||
Qy     104 AENAYEYFTKIATSLFESGN--WGRVVALLGFGYRLALHVYQHGLTGFLGQVTRFVVDFM 161

Db      95 lrerllgwiqdqggwdgllsyfgtptwqtvtifvagvltasltiwk 140
           |:  :  || ::||| : |:    |    | : :: || : : :::
Qy     162 LHHCIARWIAQRGGWVAALNLGNGPI-LNVLVVLGVVLLGQFVVRR 206


RESULT   15
ID   HSA1_MOUSE     STANDARD;      PRT;   172 AA.
AC   Q07440;
DT   01-FEB-1995 (REL. 31, CREATED)
DT   01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)
DT   01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)
DE   HEMOPOIETIC-SPECIFIC EARLY RESPONSE PROTEIN (A1 PROTEIN).
GN   A1.
OS   MUS MUSCULUS (MOUSE).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; RODENTIA.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CBA/J; TISSUE=BONE MARROW;
RX   MEDLINE; 93346743.
RA   LIN E.Y., ORLOFSKY A., BERGER M.S., PRYSTOWSKY M.B.;
RL   J. IMMUNOL. 151:1979-1988(1993).
CC   -!- FUNCTION: MAY FUNCTION IN THE RESPONSE OF HEMOPOIETIC CELLS TO
CC       EXTERNAL SIGNALS.
CC   -!- INDUCTION: BY GRANULOCYTE-MACROPHAGE COLONY-STIMULATING FACTOR
CC       AND LPS IN MACROPHAGES.
CC   -!- SUBCELLULAR LOCATION: INTRACELLULAR.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN HEMOPOIETIC TISSUES, INCLUDING
CC       BONE MARROW, SPLEEN AND THYMUS.
CC   -!- SIMILARITY: BELONGS TO THE BCL-2 FAMILY.
DR   EMBL; L16462; G293274; -.
DR   PROSITE; PS01080; BCL2.
KW   APOPTOSIS.
FT   DOMAIN       24     33       ALA/PRO-RICH.
SQ   SEQUENCE   172 AA;  19914 MW;  FFD38D6F CRC32;

  Query Match           10.2%;  Score 158;  DB 5;  Length 172;
  Best Local Similarity  28.1%;  Pred. No. 1.01e-09;
  Matches   32;  Conservative   29;  Mismatches  46;  Indels   7;  Gaps   6;

Db      28 sapsqacrvlqrvafsvqkeveknlksylddfhvesidtariifnqvmekefedgiinwg 87
           |: :|  | |  ::  : :  : :: : |: ::  :  : |   |  :    || |  |||
Qy      69 STMGQVGRQLAIIGDDINRRYDSEFQTMLQHLQ-PTAENAYEYFTKIATSLFESG--NWG 125

Db      88 rivtifafggvllkklpqeqialdvcaykqvssfvaefimnn-tgewirqnggw 140
           |:|::::||   |  |   | :|      ||: || :|::::  : || | |||
Qy     126 RVVALLGFG-YRLA-LHVYQHGLTG-FLGQVTRFVVDFMLHHCIARWIAQRGGW 176
```

Search completed: Wed Aug 20 11:14:06 1997
Job time : 26 secs.